Evaluation of Apolipoprotein A1 and Posttranslationally Modified Forms of Transthyretin as Biomarkers for Ovarian Cancer Detection in an Independent Study Population

Lee E. Moore,1 Eric T. Fung,2 Marielena McGuire,2 Charles C. Rabkin,1 Annette Molinaro,1 Zheng Wang,2 Fujun Zhang,2 Jing Wang,2 Christine Yip,2 Xiao-Ying Meng,2 and Ruth M. Pfeiffer1

1Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Department of Health and Human Services, Bethesda, Maryland and 2Ciphergen Biosystems, Inc., Fremont, California

Abstract

Background: Although overall 5-year survival rates for ovarian cancer are poor (10-30%), stage I/IIa patients have a 95% 5-year survival. New biomarkers that improve the diagnostic performance of existing tumor markers are critically needed. A previous study by Zhang et al. reported identification and validation of three biomarkers using proteomic profiling that together improved early-stage ovarian cancer detection.

Methods: To evaluate these markers in an independent study population, postdiagnostic/pretreatment serum samples were collected from women hospitalized at the Mayo Clinic from 1980 to 1989 as part of the National Cancer Institute Immunodiagnostic Serum Bank. Sera from 42 women with ovarian cancer, 65 with benign tumors, and 76 with digestive diseases were included in this study. Levels of various posttranslationally forms of transthyretin and apolipoprotein A1 were measured in addition to CA125.

Results: Mean levels of five of the six forms of transthyretin were significantly lower in cases than in controls. The specificity of a model including transthyretin and apolipoprotein A1 alone was high [96.5%; 95% confidence interval (95% CI), 91.9-98.8%] but sensitivity was low (52.4%; 95% CI, 36.4-68.0%). A class prediction algorithm using all seven markers, CA125, and age maintained high specificity (94.3%; 95% CI, 89.1-97.5%) but had higher sensitivity (78.6%; 95% CI, 63.2-89.7%).

Conclusions: We were able to replicate the findings reported by Zhang et al. in an independently conducted blinded study. These results provide some evidence that including age of patient and these markers in a model may improve specificity, especially when CA125 levels are ≥35 units/mL. Influences of sample handling, subject characteristics, and other covariates on biomarker levels require further consideration in discovery and replication or validation studies. (Cancer Epidemiol Biomarkers Prev 2006;15(9):1641–6)

Introduction

Recent advances in proteomic profiling technologies have made it possible to associate changes in protein expression with disease conditions, allowing the identification of biomarkers that can be combined to generate a multimarker panel to improve disease diagnosis. In particular, there have been several attempts to use serum proteins for the detection of early-stage ovarian cancer (1-4). In ovarian cancer, more than two thirds of cases are detected at an advanced stage, resulting in poor overall 5-year survival rates of 10% to 30% (5). This is in stark contrast to stage I/IIa patients with 95% 5-year survival (5). Longitudinal studies are under way in Europe, Japan, and the United States to evaluate screening strategies using CA125 and/or transvaginal sonography and their effect on overall cancer mortality. Although it is not known whether a survival benefit will be observed among patients diagnosed early through a screening regimen, the existing serum markers such as CA125, CA 72-4, and macrophage colony stimulating factor do not have adequate sensitivity or specificity to be used as screening tools (6). Proteomic technologies have been used to search for new biomarkers that may improve the diagnostic performance of existing markers.

In one study, Zhang et al. (2) reported using surface- enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF-MS) to identify three biomarkers that simultaneously improved the detection of early-stage ovarian cancer, in particular test specificity. Subsequently, quantitative chromatographic assays were developed for biomarkers identified at mass to charge ratios (m/z) 28,043 (apolipoprotein A1), 12,828 (a truncated form of transthyretin), and 3,272 (a fragment of inter-α-trypsin inhibitor IV). These three markers were found to differentiate ovarian cancer cases from healthy women with higher accuracy than CA125 alone. Applied to an independent validation sample set of sera from early-stage ovarian cancer cases and healthy controls, the sensitivity of a multivariable model combining the three biomarkers and CA125 at a matched specificity of 97% [95% confidence interval (95% CI), 89-100%] was 74% (95% CI, 52-90%), a small improvement over that of CA125 alone (65%; 95% CI, 43-84%; ref. 2). When compared at a fixed sensitivity of 83% (95% CI, 61-95%), the specificity of the model including CA125 and the proteomic biomarkers (94%; 95% CI, 85-98%) was significantly better than that of CA125 alone (52%; 95% CI, 39-65%).

To evaluate the discriminatory power of these markers, we measured them in an independent study population using a newly developed chromatographic SELDI-TOF-MS–based assay for quantification. We analyzed postdiagnostic/
pretreatment serum samples collected from women hospital-
ized with ovarian cancer, benign ovarian tumors, and digestive
disorders (hernias and gallstones) at the Mayo Clinic and
stored at the National Cancer Institute Immunodiagnostic
Serum Bank for apolipoprotein A1 and posttranslationally
modified forms of transthyretin because these biomarkers
could significantly discriminate ovarian cancer cases from
controls (2). Cautioned by a number of recent articles that have
called into question the reproducibility and relevance of
reported proteomic biomarkers in cancer detection (6-9), and
the absence of positive validation studies in the literature, we
aimed to carefully address several previously raised points of
criticism in our analysis, including sources of marker
variability among noncancer controls.

Materials and Methods

Patient Population. Serum samples (n = 238) were selected
from subjects whose blood was collected at the Mayo Clinic
between 1980 and 1989 (10). Once collected, samples were
shipped to the National Cancer Institute Immunodiagnosis
Serum Bank (Rockville, MD) and stored at -70 °C to -76 °C
until use. The National Cancer Institute Immunodiagnosis
Serum Bank contains -70 °C cryopreserved sera collected
between 1980 and 1989 from Mayo Clinic patients diagnosed
with a wide variety of malignant, benign, and nonneoplastic
conditions (10). For the present study, sera were selected from
each available sample from women with malignant (n = 45) or
benign (n = 71) ovarian tumors, and from 122 female controls
with abdominal hernias (ICD-9CM codes 553.1-553.3) or
gallstones (ICD-9CM codes 574.1-574.4), frequency matched
for age to cases. Information available on all subjects included
age of patient at diagnosis, smoking status (never, past current,
number of packs per day, years of smoking), ICD code for
disease as well as the draw date and number of freeze-thaws
for serum samples. Histologic subtype was known for all
cancer cases, and from 122 female controls with abdominal
hernias (ICD-9CM codes 553.1-553.3) or gallstones (ICD-9CM
codes 574.1-574.4), frequency matched for age to cases. Infor-
mation available on all subjects included age of patient at
diagnosis, smoking status (never, past current, number of
packs per day, years of smoking), ICD code for
disease as well as the draw date and number of freeze-thaws
for serum samples. Histologic subtype was known for all
malignant and benign tumors. Stage and grade were provided
for all cases. Six individuals whose serum samples were
previously thawed and refrozen ≥1 times were excluded from
analyses. Two cases were excluded for divergent histology
(one diagnosed with mesothelioma and another with signet
ring cell carcinoma) and one benign tumor was excluded for
early age at diagnosis (3 years). Therefore, the final analyses
included 42 ovarian cancer cases, 65 women with benign
tumors, and 122 noncancer control digestive diseases.

Laboratory Methods

Materials. Sinapinic acid (5-mg vial; Ciphergen, Fremont,
CA); sample denaturation buffer (Ciphergen); OMAC-30 array
(Ciphergen); Bioprocessor (Ciphergen); Q10 ProteinChip
Array (Ciphergen), human prealbumin, purified (Biosdesign
International Saco, ME), and apolipoprotein A1 calibrators
(K-ASSAY, Kamiya Biomedical, Seattle, WA).

Assays for Transthyretin and Apolipoprotein A1. To quan-
titatively measure and compare apolipoprotein A1 and posttrans-
lationally modified forms of transthyretin concentrations in
patient sera, a SELDI-TOF-MS Protein Chip array chromato-
graphic assay was developed for each marker. The following
procedures were done on a Tecan Aquarius-96 robotic
workstation. Assays were run in triplicate. For apolipoprotein
A1, IMAC ProteinChip Arrays were precharged with 50 μL
of 50 mmol/L CuSO4 for 10 minutes, was washed four times
with deionized water, and then equilibrated with IMAC binding/
washing buffer [50 mmol/L sodium phosphate, 250 mmol/L
NaCl (pH 6.0)], twice each for 5 minutes. Five microliters
of serum sample were denatured with 7.5 μL of sample
denaturation buffer [9 mol/L urea, 2% CHAPS 50 mmol/L
Tris-HCl (pH 9.0)] for 20 minutes on a shaker at room
temperature. The spots were then rinsed twice with 150 μL of the
binding/washing buffer, pipetting up and down 10 times for
each wash. The spots were then rinsed twice with 150 μL of
water. Excess water was aspirated and the spots allowed to
air-dry for 10 minutes. To each spot, 1 μL of sinapinic acid
matrix dissolved in 50% acetonitrile/0.5% trifluoroacetic acid
in water at a concentration of 12.5 mg/mL was deposited. After
allowing the spots to air-dry for 10 minutes, matrix was added
again. The SELDI-based chromatographic assay for tran-
thyretin was done using the anion exchange Q10 ProteinChip
Array as previously described (11). A set of transthyretin
calibrators (purified human prealbumin was reconstituted in
binding/washing buffer) and a set of apolipoprotein A1
calibrators were used to monitor assay performance and assay
linearity. Calibrants (representing serial dilutions of transthy-
retin and apolipoprotein A1) were treated exactly as serum
samples. In the experimental runs, each cassette incorporated
one set of calibrators with the remaining samples.

Data Acquisition. The arrays were read in a PCS4000
ProteinChip Reader, a time-lag focusing, linear laser deser-
ption/ionization-time-of-flight mass spectrometer. The instru-
ments were internally calibrated on a daily basis. All spectra
were acquired in the positive-ion mode. Time-lag focus mass
was set at 14,000 Da for transthyretin and 28,000 Da for
apolipoprotein A1. Sampling rate was set at 800 MHz. Ions
were extracted using 3.4-kV ion extraction pulse and acceler-
ated to final velocity using 25-kV acceleration potential. The
system employed a pulsed nitrogen laser at repetition rate of

Figure 1. A. Posttranslationally modified transthyretin peaks
quantitated on a Q10 ProteinChip Array. B. Apolipoprotein A1 peaks
quantitated on an IMAC ProteinChip Array.
Table 1. Description of subjects included in ovarian cancer early detection study

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Controls (N = 76)</th>
<th>Benign tumor (N = 65)</th>
<th>Cases (N = 42)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y), mean (range)</td>
<td>59 (19-88)</td>
<td>41 (15-74)</td>
<td>61 (21-78)</td>
</tr>
<tr>
<td>Smoking status, N (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>48 (63%)</td>
<td>37 (57%)</td>
<td>29 (69%)</td>
</tr>
<tr>
<td>Past</td>
<td>13 (17%)</td>
<td>11 (17%)</td>
<td>11 (26%)</td>
</tr>
<tr>
<td>Current</td>
<td>10 (13%)</td>
<td>17 (26%)</td>
<td>2 (5%)</td>
</tr>
<tr>
<td>Packs/d (mean ± SD)</td>
<td>1.3 ± 2.9</td>
<td>1.0 ± 2.1</td>
<td>0.7 ± 2.1</td>
</tr>
<tr>
<td>CA125 (units/mL), median (interquartile range)</td>
<td>15.1 (9.6-113.7)</td>
<td>18.3 (20.7-216.5)</td>
<td>468.0 (793.8-17,154)</td>
</tr>
<tr>
<td>CA125 &gt;35 units/mL, N (%)</td>
<td>7 (9.21%)</td>
<td>16 (24.62%)</td>
<td>37 (88.1%)</td>
</tr>
<tr>
<td>Years in storage (mean ± SD)</td>
<td>20.0 ± 1.2</td>
<td>17.2 ± 1.6</td>
<td>16.8 ± 1.3</td>
</tr>
<tr>
<td>Tumor grade, N (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well differentiated</td>
<td></td>
<td></td>
<td>6 (14%)</td>
</tr>
<tr>
<td>Moderately differentiated</td>
<td></td>
<td></td>
<td>6 (14%)</td>
</tr>
<tr>
<td>Poorly/undifferentiated</td>
<td></td>
<td></td>
<td>29 (69%)</td>
</tr>
<tr>
<td>Unknown</td>
<td></td>
<td></td>
<td>1 (2%)</td>
</tr>
<tr>
<td>Tumor stage, N (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I/II</td>
<td>21 (9.1%)</td>
<td>9 (21%)</td>
<td>33 (79%)</td>
</tr>
<tr>
<td>III/IV</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

20 Hz. Laser pulse energy of 1,500 to 2,000 nJ was delivered into a 100-μm diameter area, and this illuminated area was rastered across a 2-mm diameter sample spot. An automated analytic protocol was used to control the data acquisition process in most of the sample analysis. Each spectrum was an average of at least 1,000 laser shots and externally calibrated against a mixture of known peptides or proteins.

Data Processing. Raw data obtained from the PCS4000 ProteinChip Reader were first smoothed by a fixed-width moving average filter of 25 data points, and then a convex hull baseline subtraction algorithm was applied to the smoothed data. Data were then internally normalized using total ion current with the Ciphergen Express 3.0 software. Six peaks corresponding to transthyretin biomarkers, including a truncated form (T1; m/z 12,852), unmodified (T2; m/z 13,773), and four posttranslationally modified forms [sulfonated (T3; +SO2H; m/z 13,857), cysteinylated (T4; Cys; m/z 13,893), cysteinylated and gliysinylated (T5; +Cys-Gly; m/z 13,933), and glutathionylated (T6; +Glut; m/z 14,111)], were manually labeled and their intensity recorded from the Q10 ProteinChip Array data while blinded to disease status (Fig. 1A). A peak corresponding to serum apolipoprotein A1 located at m/z 28,107 (A1) was manually selected and its intensity recorded from the IMAC array data (Fig. 1B).

Immunoaassay Protein Analyses. CA125 levels (units/mL) were obtained using an Elecsys 1010 immunoaassay analyzer (Roche Diagnostics, Indianapolis, IN). Transthyretin levels (mg/mL) were obtained using an immunoprecipitation procedure (imunoturbidimetric assay; Pacific Biometrics, Seattle, WA). Samples were mixed with a polymeric enhancer and antise-

Downloaded from cebp.aacrjournals.org on August 27, 2017. © 2006 American Association for Cancer Research.
due to frequency matching on age. Benign ovarian tumor patients were younger than cases or digestive disease controls ($P < 0.0001$). The majority of serum samples from women with benign tumors and ovarian malignancies were collected between 1986 and 1989 (89% and 95%, respectively) whereas 66% of digestive disease control samples were collected before 1986 (data not shown). More women with benign tumors reported currently and ever smoking cigarettes. Smoking prevalence did not statistically differ between cancer cases, benign tumor, and digestive disease controls. Only CA125 and T1 were higher in cases compared with controls. Protein markers T2 to T6 and apolipoprotein A1 were lower in cases. The mean levels of CA125, total transthyretin, and protein markers T1 to T6 were significantly different between case, benign tumor, and control groups in a model adjusted for age quartile and storage time indicator. Protein markers T1 to T3 decreased and apolipoprotein A1 increased significantly with age of patient (data not shown). Serum CA125 and transthyretin protein levels were not associated with smoking status, age of patient, or sample storage time (data not shown).

In Table 5, the cross-validated sensitivity and specificity estimates for three prediction models are presented for the $K$-nearest neighbor algorithm with $K = 2$. Model 1 included age of patient in quartiles and an indicator variable for CA125 (<35 versus ≥35 units/mL). Model 2 included age of patient in quartiles and protein biomarker levels measured using the chromatographic assay. Model 3 was the most comprehensive and included age in quartiles, the CA125 indicator variable, and all protein biomarkers measured with the chromatographic assay. Sensitivity and specificity estimates to predict cancer/noncancer status were done by combining all noncancer subgroups. Models 1 and 3 were the most sensitive to predict cancer/noncancer status [73.8% (95% CI, 58.0-86.1%) and 78.6% (95% CI, 63.2-89.7%), respectively]. Model 2 had the

**Table 2. Histologic subtypes of ovarian benign disease and cancer**

<table>
<thead>
<tr>
<th>Histologic subtypes of benign tumors</th>
<th>N (%)</th>
<th>Histologic subtypes of ovarian cancer cases</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenofibroma, NOS</td>
<td>5 (7.7)</td>
<td>Adenocarcinoma, NOS</td>
<td>7 (16.7)</td>
</tr>
<tr>
<td>Brenner tumor, NOS</td>
<td>1 (1.5)</td>
<td>Adenosquamous carcinoma</td>
<td>1 (2.4)</td>
</tr>
<tr>
<td>Cystadenoma, NOS</td>
<td>1 (1.5)</td>
<td>Brenner Tumor malignant</td>
<td>1 (2.4)</td>
</tr>
<tr>
<td>Deroild endo, NOS</td>
<td>9 (13.8)</td>
<td>Carcinoma, NOS</td>
<td>2 (4.8)</td>
</tr>
<tr>
<td>Fibroma, NOS</td>
<td>5 (7.7)</td>
<td>Carcinoma, anaplastic, NOS</td>
<td>1 (2.4)</td>
</tr>
<tr>
<td>Leiomyoma, NOS</td>
<td>1 (1.5)</td>
<td>Clear cell adeno carcinoma, NOS</td>
<td>1 (2.4)</td>
</tr>
<tr>
<td>Mucinous cystadenoma, NOS</td>
<td>9 (13.8)</td>
<td>Endometrioid carcinoma</td>
<td>3 (7.1)</td>
</tr>
<tr>
<td>Papillary cystadenoma, NOS</td>
<td>1 (1.5)</td>
<td>Mucinous cystadenocarcinoma, NOS</td>
<td>3 (7.1)</td>
</tr>
<tr>
<td>Serous adenofibroma</td>
<td>5 (7.7)</td>
<td>Mullerian mixed tumor</td>
<td>1 (2.4)</td>
</tr>
<tr>
<td>Serous cystadenoma, NOS</td>
<td>8 (12.5)</td>
<td>Papillary adenocarcinoma, NOS</td>
<td>3 (7.1)</td>
</tr>
<tr>
<td>Terotoma, benign</td>
<td>16 (24.6)</td>
<td>Papillary mucinous cystadenocarcinoma</td>
<td>2 (4.8)</td>
</tr>
<tr>
<td>Thecoma</td>
<td>3 (4.6)</td>
<td>Papillary serous cystadenocarcinoma</td>
<td>10 (23.8)</td>
</tr>
<tr>
<td>Missing</td>
<td>1 (1.5)</td>
<td>Serous cystadenocarcinoma, NOS</td>
<td>7 (16.7)</td>
</tr>
<tr>
<td>Total</td>
<td>65 (100)</td>
<td>Total</td>
<td>42 (100)</td>
</tr>
</tbody>
</table>

In Table 3, Spearman correlation coefficients between markers quantified with the chromatographic and immunoassay-based assays are presented separately for ovarian cancer cases, benign tumor, and digestive disease controls. In each group, the rank of the transthyretin protein levels measured with the immunoassay was correlated with those of the posttranslationally modified forms of transthyretin (T2-T6) and apolipoprotein A1 biomarkers, with the exception of truncated transthyretin (T1) and CA125. CA125 levels were correlated with most biomarkers among cases, with the exception of markers T1 (correlation = 0.25, $P = 0.11$) and T3 (correlation = −0.36, $P = 0.18$). CA125 was not associated with levels of any other biomarker among benign tumor or digestive disease controls.

In Table 4, the unadjusted least squares mean and SE estimates of serum protein levels are given separately for cancer cases, benign tumor, and digestive disease controls. Only CA125 and T1 were higher in cases compared with controls. Protein markers T2 to T6 and apolipoprotein A1 were lower in cases. The mean levels of CA125, total transthyretin, and protein markers T1 to T6 were significantly different between case, benign tumor, and control groups in a model adjusted for age quartile and storage time indicator. Protein markers T1 to T3 decreased and apolipoprotein A1 increased significantly with age of patient (data not shown). Serum CA125 and transthyretin protein levels were not associated with smoking status, age of patient, or sample storage time (data not shown).

In Table 5, the cross-validated sensitivity and specificity estimates for three prediction models are presented for the $K$-nearest neighbor algorithm with $K = 2$. Model 1 included age of patient in quartiles and an indicator variable for CA125 (<35 versus ≥35 units/mL). Model 2 included age of patient in quartiles and protein biomarker levels measured using the chromatographic assay. Model 3 was the most comprehensive and included age in quartiles, the CA125 indicator variable, and all protein biomarkers measured with the chromatographic assay. Sensitivity and specificity estimates to predict cancer/noncancer status were done by combining all noncancer subgroups. Models 1 and 3 were the most sensitive to predict cancer/noncancer status [73.8% (95% CI, 58.0-86.1%) and 78.6% (95% CI, 63.2-89.7%), respectively]. Model 2 had the

**Table 3. Correlation of transthyretin and CA125 immunoassay levels with apolipoprotein A1 and posttranslationally modified forms of transthyretin by disease status**

<table>
<thead>
<tr>
<th>CA125 level (units/mL), correlation coefficient* ($P$)</th>
<th>Transthyretin level (mg/mL), correlation coefficient* ($P$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>Benign</td>
</tr>
<tr>
<td>CA125 (units/mL)</td>
<td>1.00</td>
</tr>
<tr>
<td>Transthyretin (mg/mL)</td>
<td>-0.09 (0.47)</td>
</tr>
<tr>
<td>T1-m/z 12,852</td>
<td>-0.14 (0.22)</td>
</tr>
<tr>
<td>T2-m/z 13,773</td>
<td>-0.17 (0.23)</td>
</tr>
<tr>
<td>T3-m/z 13,857</td>
<td>-0.14 (0.23)</td>
</tr>
<tr>
<td>T4-m/z 13,893</td>
<td>0.03 (0.81)</td>
</tr>
<tr>
<td>T5-m/z 13,933</td>
<td>0.03 (0.77)</td>
</tr>
<tr>
<td>T6-m/z 14,111</td>
<td>0.09 (0.47)</td>
</tr>
<tr>
<td>A1-m/z 28,107</td>
<td>-0.08 (0.49)</td>
</tr>
</tbody>
</table>

*Spearman rank correlation coefficient.  
*Immunooassay measurement.
Table 4. Least squares mean estimates for patient status by cancer marker

<table>
<thead>
<tr>
<th>Biomarkers</th>
<th>Control</th>
<th>Unadjusted model mean (SD)</th>
<th>P-adjusted* model</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA125 (units/mL)</td>
<td>20.5 (19.3)</td>
<td>29.9 (32.7)</td>
<td>1,563.9 (3,350.0)</td>
</tr>
<tr>
<td>Transferrin (mg/mL)</td>
<td>294.8 (78.6)</td>
<td>287.1 (73.8)</td>
<td>221.9 (89.1)</td>
</tr>
<tr>
<td>Transhyretin (mL)</td>
<td>2.8 (0.8)</td>
<td>2.8 (0.6)</td>
<td>3.6 (1.5)</td>
</tr>
<tr>
<td>T1-12,552</td>
<td>9.3 (4.3)</td>
<td>9.6 (4.2)</td>
<td>7.3 (4.3)</td>
</tr>
<tr>
<td>T2-13,773</td>
<td>5.6 (1.8)</td>
<td>6.1 (1.8)</td>
<td>4.5 (2.2)</td>
</tr>
<tr>
<td>T3-13,857</td>
<td>18.8 (6.2)</td>
<td>17.0 (6.2)</td>
<td>15.5 (8.3)</td>
</tr>
<tr>
<td>T4-13,893</td>
<td>6.8 (2.3)</td>
<td>7.2 (2.9)</td>
<td>5.4 (2.6)</td>
</tr>
<tr>
<td>T5-13,933</td>
<td>4.4 (0.9)</td>
<td>4.5 (1.0)</td>
<td>3.7 (1.5)</td>
</tr>
<tr>
<td>Apolipoprotein A1 (mL)</td>
<td>9.3 (4.3)</td>
<td>8.2 (3.3)</td>
<td>8.3 (3.1)</td>
</tr>
</tbody>
</table>

*Significance of group differences in model adjusted for age quartile and storage time indicator variable.

Table 5. Cross-validated sensitivity and specificity estimates for various prediction models using the K-nearest neighbor algorithm with K = 2 (proteomic markers, when included into a model, were log transformed)

<table>
<thead>
<tr>
<th>Sensitivity and specificity (95% CI) to discriminate cancer from noncancer</th>
<th>Sensitivity and specificity (95% CI) by subgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sensitivity</td>
<td>Total specificity</td>
</tr>
<tr>
<td><strong>Model 1</strong></td>
<td></td>
</tr>
<tr>
<td>31/42 (73.8%; 58.0-86.1%)</td>
<td>133/141 (94.3%; 89.1-97.5%)</td>
</tr>
<tr>
<td><strong>Model 2</strong></td>
<td></td>
</tr>
<tr>
<td>22/42 (52.4%; 36.4-68.0%)</td>
<td>136/141 (96.5%; 91.9-98.8%)</td>
</tr>
<tr>
<td><strong>Model 3</strong></td>
<td></td>
</tr>
<tr>
<td>33/42 (78.6%; 65.2-89.7%)</td>
<td>133/141 (94.3%; 89.1-97.5%)</td>
</tr>
</tbody>
</table>


*Exact binomial confidence intervals.

**Model 1**: CA125 ≥35, age ≥35, and age in quartiles.

**Model 2**: All proteomic markers (T1-T6, ApoA1) and age in quartiles.

**Model 3**: CA125 ≥35, ≥35, all proteomic markers (T1-T6, ApoA1), and age in quartiles.
“omic” technologies that are potentially able to differentiate healthy controls from ovarian cancer patients are also found in controls, associated with other disease states, and are considered to be markers of the host metabolic response to cancer rather than specific to the cancer itself (2, 4, 11, 22-24). The major challenge of developing an ovarian cancer detection test is that it must be highly specific to avoid detection of numerous false positives. This requirement and the fact that many markers are not specific to ovarian cancer underscore the need to understand how they are associated with sample processing and handling and other patient characteristics, in addition to their significance in pathogenesis. In this study, we examined associations between all markers and available information on subject age at diagnosis, smoking status, ICD disease code, as well as draw date and number of freeze thaws per sample. It was necessary to include age quartile in all models because it was positively associated with apolipoprotein A1, inversely associated with T1 to T3, and because the women diagnosed with benign cysts were younger than cancer cases or digestive disease controls. By eliminating age from the models, the specificity of all three models to detect controls with benign cysts would be lower (Table 5). When used alone, apolipoprotein A1 and transthyretin markers were useful to detect controls with CA125 levels ≥35 units/mL but poor at identifying cases with late-stage cancers. The addition CA125 as a dichotomous variable to the model increased sensitivity to detect high-stage cancers; however, the added specificity was lost.

This is one of the first studies to independently evaluate findings previously reported in ovarian cancer early detection using the SELDI-TOF-MS platform. Two recent editorials pointed to the important role of assessing potential sources of bias in design and conduct and also the effect of chance on data interpretations (7, 25). Several important lessons have been learned through the design and analysis of this study that might be considered in future investigations to reduce ongoing concerns over data validity and improve chances of reproducibility: (a) Convenience samples should be avoided for biomarker discovery or validation until the relationship between the biomarker and biases in sample handling, processing, storage, and common confounding variables is understood. Ignoring such factors in study design and analysis will decrease any chance of future validation. In contrast to common practice in early phases of biomarker identification and validation, we promote using prospectively collected samples and including comparison groups that are frequency matched to cases and well characterized for possible confounders. Such samples are not generally easy to obtain. (b) Inclusion of a clinically relevant comparison group, in addition to healthy population controls, may be needed to improve understanding of the biomarker performance in future testing populations. (c) When relationships between biomarkers and potential confounders are not well understood, it may be advisable to restrict the samples rather than attempt to eliminate biases by statistical modeling. In biomarker discovery studies, many questions can be addressed in small carefully designed studies to minimize possible sources of variability. Once biomarkers are identified, control samples can be used to identify relationships with possible confounding variables. Whereas ideally one restricts samples in the design phase, sometimes it is necessary to contend with them in the analysis. The major factor biasing our results was length of sample storage time. Because the digestive disease control serum samples were stored longer than those from ovarian cancer cases or benign ovarian tumor controls, the protein levels of two markers would have decreased with storage time, biasing the comparisons between cases and digestive disease controls toward the null. We thus restricted the analysis to samples collected 1983 or later, eliminating the storage artifact for all markers except T4. Analyses comparing benign ovarian
tumor and cancer cases were not susceptible to this bias because samples from both groups were collected and stored over a similar time period.

In conclusion, the specificity of a model that included measurements of posttranslationally modified forms of transthyretin, apolipoprotein A1, and age of patient was as high as models that included CA125 and age but sensitivity was lower. When used alone, these markers improved detection of controls with CA125 levels ≥35 units/mL but lost sensitivity to identify late-stage cases. In general, influences of sample storage conditions, subject characteristics, and other covariates on biomarker levels require further consideration in discovery and replication studies.

References


Evaluation of Apolipoprotein A1 and Posttranslationally Modified Forms of Transthyretin as Biomarkers for Ovarian Cancer Detection in an Independent Study Population


Updated version
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/15/9/1641

Cited articles
This article cites 20 articles, 5 of which you can access for free at:
http://cebp.aacrjournals.org/content/15/9/1641.full#ref-list-1

Citing articles
This article has been cited by 4 HighWire-hosted articles. Access the articles at:
http://cebp.aacrjournals.org/content/15/9/1641.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.