

Cancer Prevention Research



Ovarian Cancer Biomarker Performance in Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial Specimens

Daniel W. Cramer, Robert C. Bast, Jr., Christine D. Berg, et al.

Cancer Prev Res 2011;4:365-374. Published online March 2, 2011.

Updated Version

Access the most recent version of this article at:
doi:[10.1158/1940-6207.CAPR-10-0195](https://doi.org/10.1158/1940-6207.CAPR-10-0195)

Supplementary Material

Access the most recent supplemental material at:
<http://cancerpreventionresearch.aacrjournals.org/content/suppl/2011/03/01/1940-6207.CAPR-10-0195.DC1.html>

Cited Articles

This article cites 25 articles, 9 of which you can access for free at:
<http://cancerpreventionresearch.aacrjournals.org/content/4/3/365.full.html#ref-list-1>

Citing Articles

This article has been cited by 6 HighWire-hosted articles. Access the articles at:
<http://cancerpreventionresearch.aacrjournals.org/content/4/3/365.full.html#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.

Research Article

Ovarian Cancer Biomarker Performance in Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial Specimens

Daniel W. Cramer¹, Robert C. Bast Jr.², Christine D. Berg³, Eleftherios P. Diamandis⁴, Andrew K. Godwin⁵, Patricia Hartge⁶, Anna E. Lokshin⁷, Karen H. Lu², Martin W. McIntosh⁸, Gil Mor⁹, Christos Patriotis¹⁰, Paul F. Pinsky³, Mark D. Thornquist⁸, Nathalie Scholler¹¹, Steven J. Skates¹², Patrick M. Sluss¹², Sudhir Srivastava¹⁰, David C. Ward¹³, Zhen Zhang¹⁴, Claire S. Zhu³, and Nicole Urban⁸

Abstract

Establishing a cancer screening biomarker's intended performance requires "phase III" specimens obtained in asymptomatic individuals before clinical diagnosis rather than "phase II" specimens obtained from symptomatic individuals at diagnosis. We used specimens from the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial to evaluate ovarian cancer biomarkers previously assessed in phase II sets. Phase II specimens from 180 ovarian cancer cases and 660 benign disease or general population controls were assembled from four Early Detection Research Network or Ovarian Cancer Specialized Program of Research Excellence sites and used to rank 49 biomarkers. Thirty-five markers, including 6 additional markers from a fifth site, were then evaluated in PLCO proximate specimens from 118 women with ovarian cancer and 474 matched controls. Top markers in phase II specimens included CA125, HE4, transthyretin, CA15.3, and CA72.4 with sensitivity at 95% specificity ranging from 0.73 to 0.40. Except for transthyretin, these markers had similar or better sensitivity when moving to phase III specimens that had been drawn within 6 months of the clinical diagnosis. Performance of all markers declined in phase III specimens more remote than 6 months from diagnosis. Despite many promising new markers for ovarian cancer, CA125 remains the single-best biomarker in the phase II and phase III specimens tested in this study. *Cancer Prev Res*; 4(3): 365–74. ©2011 AACR.

Introduction

Ovarian cancer generally presents in advanced stages with a high case/fatality ratio but has favorable survival if diagnosed earlier. However, no markers are approved for

early detection and only Cancer Antigen 125 (CA125) and human epididymis protein 4 (HE4) for monitoring disease. Many potential biomarkers for ovarian cancer have been studied (1), including panels of markers reported to have sensitivity and specificity superior to CA125 (2–5) and one panel which received FDA approval for preoperative assessment of ovarian cancer risk in women with an ovarian mass (6). Performance of markers has generally been assessed in cases at diagnosis; only CA125 has been tested in an actual screening trial in asymptomatic subjects (7).

The Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Study is a randomized trial evaluating screening modalities for these 4 cancers to determine their effect on cancer mortality (8–12). Women assigned to the ovarian screening arm received annual transvaginal ultrasound (TVU) and CA125 testing. Sera from this type of study, collected and banked prior to clinical diagnosis, have been termed "phase III" specimens as opposed to sera obtained at diagnosis termed "phase II" specimens (13). The PLCO makes specimens available to the scientific community for assessing screening biomarkers; and, in 2005 and 2006, approved distribution of 5 identical specimen sets to investigators on the basis of the biomarker performance in phase II studies to be tested in PLCO phase III samples. This article, the first of two companion reports, focuses on a comparison of individual marker performance in the

Authors' Affiliations: ¹Brigham and Women's Hospital, Boston, Massachusetts; ²MD Anderson Cancer Center, Houston, Texas; ³Early Detection Research Group, Division of Cancer Prevention, National Cancer Institute, NIH, Bethesda, Maryland; ⁴Biochemistry, Mount Sinai Hospital, Toronto, Ontario, Canada; ⁵Department of Pathology and Laboratory Medicine, University of Kansas Medical Center, Kansas City, Kansas; ⁶Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, Maryland; ⁷University of Pittsburgh, Pittsburgh, Pennsylvania; ⁸Fred Hutchinson Cancer Research Center, Seattle, Washington; ⁹Yale University, New Haven, Connecticut; ¹⁰Cancer Biomarker Research Group, Division of Cancer Prevention, National Cancer Institute, NIH, Bethesda, Maryland; ¹¹University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania; ¹²Massachusetts General Hospital, Boston, Massachusetts; ¹³Cancer Research Center, University of Honolulu, Hawaii; and ¹⁴Johns Hopkins University, Baltimore, Maryland

Note: This is the first of two companion manuscripts submitted under separate cover.

Correspondence Author: Daniel W. Cramer, Brigham and Women's Hospital, Ob/Gyn, Division of Epidemiology, 221 Longwood Ave., RFB 365, Boston, MA 02115. Phone: 617-732-4895; Fax: 617-732-4899. E-mail: drcramer@partners.org.

doi: 10.1158/1940-6207.CAPR-10-0195

©2011 American Association for Cancer Research.

phase II and phase III samples. The second report addresses validation of classification algorithms in the phase III samples.

Materials and Methods

Phase II studies

Markers to be evaluated in the PLCO specimens were informed by results from phase II studies which were either initiated specifically as a component of this study or had been previously published. Four sites initiated a phase II study under the Early Detection Research Network (EDRN) and Specialized Programs of Research Excellence (SPORE) in ovarian cancer. This "common" set consisted of serum from 160 ovarian cancer cases seen at Fred Hutchinson Cancer Research Center (FHCRC), Fox Chase Cancer Center (FCCC), MD Anderson Cancer Center (MDACC), and Brigham and Women's and Massachusetts General Hospitals (Partners Healthcare). Approximately half of these were early-stage cases. Also included was a benign disease group of 160 cases and 480 controls—the latter recruited either from the general population as controls for an ovarian cancer study or women undergoing mammographic screening. Forty paired serial specimens from women taken at annual mammographic screening at FHCRC and 84 replicates constructed from female sera pooled for laboratory standardization (ProMedDX) were included. Cases had blood drawn prior to surgery or chemotherapy, processed generally within 4 hours, and stored in aliquots at -80°C . Controls had blood drawn, processed, and stored under similar conditions at contributing sites. All subjects were accrued between 1998 and 2006 under protocols approved by each site's Institutional Review Board allowing for collection of medical information and specimens and sharing of specimens. Specimens were transferred in accordance with agreements at the shipping and receiving institutions regarding transfer of specimens without personal identifiers.

We requested sera that had not previously undergone freeze-thaw cycles. Two-milliliter aliquots of sera for each subject were assembled and sent to FCCC (A.K.G.) for aliquot construction and relabeling to blind source and status. Four aliquots were prepared and sent to FHCRC, MDACC, Partners, and UPCI. Specimens were organized to yield similar proportions of the sample types in batch sizes of 96. Coordination and data analysis of the phase II studies was done by the EDRN Coordinating Center (M.D.T.).

The markers selected for testing in the PLCO specimens from Yale University School of Medicine (YUSM) were based on two previously published phase II studies. A four marker panel was originally described in ref. 14, modified to include CA125, and developed into a kit (Cancer Panel) produced by Millipore. This five marker panel was evaluated in 156 newly diagnosed ovarian cancer patients and 362 healthy women participating in the YUSM Early Detection Clinic and achieved 95% sensitivity at 99.4% specificity (4). A second phase II study used samples from

Gynecologic Oncology Group (GOG) protocol 136 including 193 epithelial ovarian cancer patients (93 stages I and II; 100 stages III and IV) and 94 controls. Classification results for all GOG cases using the Cancer Panel were 95.3% sensitivity at 96.8% specificity and area under the curve (AUC) of 0.97 (15).

Phase III study

The phase III study was coordinated by the PLCO. Original sample collection, processing, shipping, and storage were standardized across screening sites (8–12). Bloods for research purposes were drawn and processed within 2 hours, assigned an identification number and shipped frozen to the PLCO biorepository in Frederick, MD for storage at -70°C or -157°C . By June 2006, 118 cases of invasive ovarian, primary peritoneal, and fallopian tube cancers had been confirmed with the available samples. Tumors of borderline malignancy were excluded. A single-serum sample most proximate to diagnosis was selected for cases. Controls were selected from healthy individuals who remained cancer-free and matched to cases by 5-year categories of age and calendar year at blood draw. Women with a history of cancer or oophorectomy at baseline were excluded. For each case, 8 controls were selected: 4 "general population" controls randomly selected from eligible controls; 2 with positive family history of breast or ovarian cancer; and 2 with elevated CA125 at any time. Sixty replicate pairs were randomly inserted for quality control. This report will focus on cases and 474 general population controls.

One 1.8-mL aliquot of serum for each subject was retrieved from the biorepository and centrally processed into aliquots, made side-by-side in small batches to avoid repeated freeze-thaws and minimize handling. FHCRC received 0.3 mL, MDACC 0.1 mL, Partners 0.6 mL, Pittsburgh 0.2 mL, and YUSM 0.2 mL in otherwise identical sets of aliquots labeled with a unique sample ID and shipped on dry ice to the assay sites.

Laboratory assays

Assays for the common phase II study and YUSM phase II studies, as well as the phase III PLCO study were conducted at the 5 separate laboratories by individuals blind to case or control status. Investigators at the FHCRC (N.U., N.S., M. W.M.) evaluated 5 markers using singleplex Luminex bead assays, as described for CA125 (16), HE4, and mesothelin (17). Reagents supplied by Diadexus were used for Spondin-2; Insulin-like growth factor binding protein 2 (IGFBP2) was measured using the R&D Systems Duo set (Catalog DY674E). The same assays were carried over into phase III except that a bead-based assay for Secretory Leukocyte Protease Inhibitory (SLPI) and a plate-based assay for matrix metalloproteinase 7 (MMP7; R&D Systems kit) were introduced. The SLPI assay was performed using a pool of recombinant antibodies for capture and goat anti-SLPI purified antibodies (R&D Systems) for detection. Investigators at MDACC (R.C.B., K.H.L.) and Johns Hopkins University (Z.Z.) evaluated 7 markers using a mass

spectrometer-based system for its markers in both phases further described in ref. 5. Investigators at Partners (D.W.C., P.M.S., S.J.S.) and at Mount Sinai (Ontario; EPD) evaluated 13 markers using plate-based assay for immune marker B7-H4 (Diadexus), decoy receptor 3 (DcR3; Diadexus), macrophage inhibitory cytokine 1 (MIC1; DiaDexus), Spondin-2 (DiaDexus), IGF II (Diagnostic Systems Laboratories), Mesothelin (SMRP; Fujirebio), HE4 (Fujirebio), Chitinase (YKL-40 Quidel), and kallikrein 6 (KLK6; as described in ref. 18). CA72.4 was performed by immunoradiometric assay (IRMA) (Fujirebio). Platform-based assays (Modular E170; Roche) were used for CA125, CA15.3, CA19.9, and carcinoembryonic antigen (CEA). Seven markers were carried over to phase III and these included B7-H4, CA125, CA15.3, CA19.9, CA72.4, HE4, and KLK6. The same assays were used except that a platform-based assay was available for CA72.4. Investigators at University of Pittsburgh Cancer Institute (UPCI) (A.E.L.) evaluated 34 markers in phase II using a multiplex bead assay system described in ref. 19. For phase III, the number of markers was reduced to 8 but the same technique was used. These included: CA125, HE4, Eotaxin MMP3, CA72.4, soluble vascular cell adhesion molecule 1 (sV-CAM), epidermal growth factor receptor (EGFR), and prolactin. Investigators at YUSM (G.M., D.C. W.) evaluated 6 markers in their phase II and phase III studies using a multiplex bead system described in ref. 4. These markers included CA125, macrophage inhibiting factor (MIF), leptin, prolactin, osteopontin, and insulin-like growth factor 2 (IGF2).

Statistical analysis

Assay results from the common phase II study were returned to the Data Management Committee of the EDNRN. To allow unbiased selection of top markers, results were initially reported back with marker names blinded and a "standardized value" for the marker given (20). Data were divided into training and validation sets to allow algorithms for combining markers to be developed and tested before actual marker results and names were returned. These interim analyses are not shown. On the basis of the common phase II results, the PLCO provided the phase III proximate samples in a blinded manner to the assay sites (FHCRC, MDACC, Partners, and UPCI) as well as to YUSM based on the report of their markers (4, 15). Investigators at the 5 sites completed assays on the PLCO samples and returned results (including algorithms for combining markers) to the PLCO statistician (P.F.P.). Details about the algorithms and the process for their validation are provided in the companion paper (21).

Results

Table 1 lists all markers included in the common phase II set with coefficients of variation (CV) and correlations based on the paired specimens. In general, CVs were greatest for multiplex Luminex assays, less for singleplex Lumi-

nex or plate-based assays, and least for platform-based assays of the type used in clinical laboratories. The most stable markers, based on the 40 serial paired specimens, were CA125, α -fetoprotein (AFP), CA15.3, CA19.9, CA72.4, mesothelin, and HE4. The least stable markers were prolactin, cytokeratin 19, growth hormone (GH), myeloperoxidase (MPO), TNF receptor 2 (TNFr2), and KLK6.

Characteristics of cases and controls in the common phase II and the phase III sets are shown in Table 2. Phase III subjects were older (all postmenopausal) and more likely to be Caucasian. The majority of cancers in both sets were serous type. By design, early-stage cases were oversampled in the phase II data and largely account for the high percentage of tumors of borderline malignancy.

Table 3 shows the top 20 markers found in the common phase II data ranked by sensitivity for all cases at 95% specificity and by AUCs against general population controls plus any additional markers that went on to be included in phase III. This yielded a total 32 markers shown in Table 3 of the 49 markers originally tested; all the markers not shown had sensitivities less than 0.17 and AUCs less than 0.64 in the all case columns. Top markers ranked by sensitivity include CA125, HE4, Transthyretin, CA15.3, CA72.4, and IGFBP2. AUCs give a slightly different ordering but the top three remain the same. Restricted to early-stage cases, several additional markers would be added including CA19.9, apolipoprotein A1, and prolactin (Table 3). Inclusion of benign disease controls lowered the sensitivities and AUCs of markers but did not substantially alter the rank order of markers (data not shown). Exclusion of premenopausal women or cases with borderline malignancies did not substantially change the all case rankings.

Turning to phase III data, results are shown for the 5 sites by timing of the specimen collection from clinical diagnosis in intervals of 6 months (Table 4). The CA125 value originally measured by the PLCO (as well as the remeasured CA125) performed best. The sensitivity (95% CI) was highest for cases whose blood was drawn within 6 months of diagnosis at 0.86 (0.76–0.97) and progressively declined for specimens more remote from diagnosis. For HE4, the second best marker, the sensitivity (95% CI), among cases 6 months or more from diagnosis, was 0.73 (0.60–0.86). Like CA125, its sensitivity declined in specimens more remote than 6 months from diagnosis. CVs for the assays in the PLCO data are shown in column 3. There were a total of 36 separate assays on 28 markers; the average CV for these assays ranged from 2% to 58%, with a median of 13%.

Table 5 compares the performance of top markers in moving from a phase II study to a phase III study. Phase II performances for some of the "standard" tumor markers, such as CA125, HE4, CA72.4, and CA15.3, were quite comparable to performances in a phase III study for specimens drawn within 6 months of the blood draw. However, this was not the case for markers such as prolactin, transthyretin, and transferrin. The performance of nearly all

Table 1. CV based on 84 replicates and Pearson correlations (*r*) based on 40 paired serial samples for assays in the common phase II set by site

Marker Name	Partners	FHCRC	MDACC	UPCI
AFP				0.15 (0.92)
Apolipoprotein A1			0.23 (0.39)	
B7-H4	0.29 (0.63)			
B2M			0.10 (0.64)	
CA125	0.10 (0.94)	0.12 (0.92)		0.61 (0.80)
CA15.3	0.12 (0.89)			0.26 (0.89)
CA19.9	0.09 (0.89)			0.30 (0.70)
CA72.4	0.19 (0.69)			0.14 (0.88)
CEA				0.11 (0.83)
Chitinase	0.17 (0.44)			
CTAP-III			0.12 (0.62)	
Cytokeratin 19				0.33 (0.01)
DcR3	0.06 (0.64)			
EGFR				0.13 (0.83)
Eotaxin				0.52 (0.34)
ErbB2				0.14 (0.76)
Follicle stimulating hormone				0.18 (0.81)
GH				0.29 (0.17)
HE4	0.06 (0.74)	0.25 (0.78)		0.23 (0.87)
Hepcidin			0.18 (0.56)	
IGF2	0.07 (0.72)			
IGFBP1				0.25 (0.61)
IGFBP2		0.16 (0.82)		
Interleukin 10				0.27 (0.64)
Interleukin 2 receptor				0.28 (0.55)
Interleukin 6				0.28 (0.72)
Interleukin 6 receptor				0.15 (0.84)
Interleukin 8				0.32 (0.52)
ITIH4			0.21 (0.59)	
KLK6	0.21 (0.32)			
KLK8				0.32 (0.52)
Luteinizing hormone				0.19 (0.84)
MIF				0.41 (0.52)
MIC1	0.41 (0.83)			
Mesothelin	0.42 (0.72)	0.16 (0.75)		0.18 (0.88)
MMP2				0.14 (0.42)
MMP3				0.15 (0.74)
MMP9				0.19 (0.33)
MPO				0.44 (0.23)
Prolactin				0.24 (0.01)
Soluble intracellular adhesion molecule 1				0.31 (0.43)
sV-CAM				0.25 (0.54)
Spondin-2	0.13 (0.64)	0.17 (0.63)		
TNF α				0.27 (0.38)
TNFr2				0.50 (0.28)
Total plasminogen activator inhibitor				0.32 (0.54)
Transthyretin			0.18 (0.39)	
Transferrin			0.17 (0.38)	
TSH				0.21 (0.77)

NOTE: All values are given as CV (*r*). Partners performed platform (Roche E170) or plate-based assays; FHCRC performed singleplex Luminex bead assays and plate-based assays; MDACC performed a mass spectrographic analysis using surface-enhanced antibody chips; UPCI performed multiplexed Luminex bead assays.

Abbreviations: B2M, β -2 microglobulin; CTAP-III, connective tissue activating protein III; ITIH4, inter- α -trypsin inhibitor heavy chain; TSH, thyroid stimulating hormone

Table 2. Characteristics of cases and controls in the common phase II set and the PLCO phase III data

	"Common" phase II data			Phase III data	
	Cases (n = 160)	BD controls (n = 160)	GP controls (n = 480)	Cases (n = 118)	Controls (n = 474)
Center					
FCCC	41 (25.6)	11 (6.7)	121 (25.2)	—	—
FHCRC	39 (24.4)	39 (24.4)	118 (24.6)	—	—
Partners	40 (25.0)	110 (68.8)	121 (25.2)	—	—
MDACC	40 (25.0)	0 (0.0)	120 (25.0)	—	—
Age					
30–55	68 (42.5)	73 (45.6)	209 (44.8)	0 (0.0)	0 (0.0)
55–59	31 (19.4)	26 (16.2)	88 (18.9)	14 (11.9)	60 (12.7)
60–64	21 (13.1)	28 (17.5)	73 (15.7)	37 (31.4)	146 (30.8)
≥65	40 (25.0)	33 (20.6)	96 (20.6)	67 (56.8)	268 (56.5)
Menopausal status					
Pre	44 (30.6)	43 (31.2)	139 (36.7)	0 (0.0)	0 (0.0)
Post	100 (69.4)	95 (68.8)	240 (63.3)	118 (100.0)	474 (100.0)
Race					
White	122 (76.2)	127 (79.4)	338 (73.2)	106 (89.8)	431 (90.9)
Nonwhite	38 (23.8)	33 (20.6)	124 (26.8)	12 (10.2)	43 (9.1)
Body mass index					
<22	21 (18.3)	24 (18.3)	41 (20.5)	22 (18.6)	66 (14.0)
22–24.9	32 (27.8)	38 (29.0)	59 (29.5)	31 (26.3)	113 (24.0)
25–29.9	23 (20.0)	38 (29.0)	53 (26.5)	37 (31.4)	178 (37.9)
≥30	39 (33.9)	31 (23.7)	47 (23.5)	28 (23.7)	113 (24.1)
Smoking					
Never	68 (59.6)	73 (55.7)	176 (55.5)	64 (54.2)	291 (61.4)
Former	29 (25.4)	40 (30.5)	112 (35.3)	46 (39.0)	149 (31.4)
Current	17 (14.9)	18 (13.7)	29 (9.2)	8 (6.8)	34 (7.2)
Case characteristics					
Histology					
Serous	88 (55.0)	—	—	68 (57.6)	—
Mucinous	14 (8.8)	—	—	2 (1.7)	—
Endometrioid	17 (10.6)	—	—	14 (11.9)	—
Clear cell	9 (5.6)	—	—	4 (3.4)	—
Squamous cell	1 (0.6)	—	—	0 (0.0)	—
Mixed epithelial	23 (14.4)	—	—	0 (0.0)	—
Undifferentiated	1 (0.6)	—	—	2 (1.7)	—
Other	5 (3.1)	—	—	3 (2.5)	—
Adenocarcinoma NOS	0 (0.0)	—	—	22 (18.6)	—
Unknown	2 (1.2)	—	—	0 (0.0)	—
Granulosa cell	0 (0.0)	—	—	3 (2.5)	—
Grade					
0	33 (20.6)	—	—	—	—
1	11 (6.9)	—	—	7 (5.9)	—
2	11 (6.9)	—	—	22 (18.6)	—
3–4	85 (53.1)	—	—	66 (55.9)	—
Unknown	20 (12.5)	—	—	23 (19.5)	—
Stage					
Early	74 (46.2)	—	—	32 (27.1)	—
Late	83 (51.9)	—	—	84 (71.2)	—
Unknown	3 (1.9)	—	—	2 (1.7)	—

Table 3. Phase II results for all cases and early-stage cases compared with general population controls by sensitivity at 95% specificity and AUC (sorted by sensitivity for all cases)

	All cases and general population controls			Early-stage cases and general population controls		
	Sensitivity (95% CI)	Rank	AUC (95% CI)	Sensitivity (95% CI)	Rank	AUC (95% CI)
CA125	0.73 (0.64–0.84)	1	0.92 (0.89–0.95)	0.56 (0.49–0.72)	1	0.89 (0.85–0.93)
HE4	0.57 (0.50–0.70)	2	0.86 (0.83–0.89)	0.29 (0.19–0.45)	5	0.75 (0.70–0.82)
Transferrin	0.47 (0.38–0.56)	3	0.80 (0.75–0.84)	0.29 (0.21–0.37)	4	0.71 (0.64–0.78)
CA15.3	0.46 (0.39–0.54)	4	0.74 (0.69–0.78)	0.33 (0.23–0.44)	2	0.63 (0.57–0.70)
CA72.4	0.40 (0.33–0.52)	5	0.75 (0.71–0.80)	0.20 (0.12–0.31)	9	0.63 (0.57–0.69)
IGFBP2	0.38 (0.30–0.45)	6	0.74 (0.69–0.79)	0.20 (0.11–0.29)	11	0.63 (0.57–0.71)
Cytokeratin 19	0.37 (0.25–0.45)	7	0.74 (0.69–0.77)	0.14 (0.03–0.22)	22	0.61 (0.55–0.68)
IGF2	0.36 (0.28–0.49)	8	0.80 (0.76–0.83)	0.15 (0.08–0.26)	17	0.69 (0.63–0.76)
KLK6	0.36 (0.26–0.43)	9	0.71 (0.65–0.76)	0.12 (0.04–0.20)	29	0.59 (0.52–0.66)
Mesothelin	0.35 (0.27–0.46)	10	0.71 (0.65–0.76)	0.12 (0.06–0.22)	27	0.60 (0.52–0.69)
B7-H4	0.35 (0.27–0.42)	11	0.67 (0.63–0.73)	0.19 (0.10–0.28)	13	0.59 (0.52–0.67)
Apolipoprotein A1	0.34 (0.26–0.43)	12	0.77 (0.73–0.80)	0.28 (0.15–0.39)	6	0.68 (0.62–0.74)
sv-CAM	0.34 (0.22–0.42)	13	0.75 (0.71–0.79)	0.27 (0.08–0.41)	8	0.67 (0.60–0.74)
Prolactin	0.34 (0.15–0.49)	14	0.77 (0.74–0.81)	0.28 (0.12–0.41)	7	0.69 (0.63–0.76)
Chitinase	0.31 (0.23–0.38)	15	0.74 (0.69–0.80)	0.13 (0.06–0.21)	24	0.63 (0.57–0.69)
Spondin-2	0.28 (0.18–0.36)	16	0.74 (0.69–0.79)	0.11 (0.03–0.18)	31	0.62 (0.55–0.68)
MIC1	0.24 (0.14–0.30)	17	0.62 (0.57–0.67)	0.16 (0.08–0.26)	15	0.54 (0.50–0.62)
CA19.9	0.23 (0.18–0.30)	18	0.61 (0.57–0.66)	0.31 (0.23–0.46)	3	0.74 (0.69–0.80)
Transferrin	0.23 (0.15–0.32)	19	0.71 (0.66–0.75)	0.19 (0.08–0.30)	14	0.65 (0.60–0.72)
IGFBP1	0.21 (0.15–0.31)	20	0.72 (0.68–0.76)	0.11 (0.04–0.19)	30	0.64 (0.57–0.70)
Hepcidin	0.21 (0.12–0.28)	21	0.62 (0.57–0.67)	0.12 (0.05–0.22)	28	0.57 (0.51–0.65)
TSH	0.19 (0.13–0.25)	22	0.60 (0.56–0.65)	0.15 (0.07–0.21)	19	0.54 (0.50–0.60)
CTAP-III	0.19 (0.12–0.24)	23	0.51 (0.50–0.57)	0.15 (0.08–0.24)	18	0.51 (0.50–0.58)
MMP3	0.17 (0.11–0.27)	24	0.64 (0.59–0.67)	0.14 (0.07–0.28)	21	0.60 (0.53–0.67)
MMP2	0.17 (0.08–0.28)	25	0.64 (0.59–0.69)	0.20 (0.08–0.32)	12	0.62 (0.55–0.69)
GH	0.16 (0.10–0.23)	26	0.67 (0.61–0.72)	0.15 (0.06–0.26)	20	0.63 (0.57–0.69)
MIF	0.15 (0.09–0.24)	28	0.66 (0.62–0.71)	0.16 (0.06–0.22)	16	0.58 (0.53–0.64)
Eotaxin	0.15 (0.07–0.23)	30	0.64 (0.59–0.69)	0.10 (0.05–0.18)	33	0.61 (0.54–0.67)
MPO	0.14 (0.09–0.22)	31	0.55 (0.50–0.61)	0.20 (0.11–0.30)	10	0.70 (0.62–0.75)
EGFR	0.10 (0.05–0.17)	35	0.64 (0.60–0.69)	0.09 (0.02–0.21)	34	0.59 (0.52–0.68)
ITIH4	0.09 (0.05–0.16)	37	0.57 (0.51–0.62)	0.04 (0.00–0.10)	48	0.51 (0.50–0.58)
B2M	0.05 (0.02–0.12)	45	0.56 (0.51–0.61)	0.05 (0.00–0.12)	47	0.50 (0.50–0.58)

Table 4. Phase III results (AUC and sensitivity at 95% specificity) by the 5 sites

Site	Marker	CV ^a	Cases diagnosed ≤6 mo after draw (n = 45)		Cases diagnosed >6–12 mo after draw (n = 22)		Cases diagnosed >12–18 mo after draw (n = 17)		Cases diagnosed >18 mo after draw (n = 34)	
			AUC	Sensitivity (95% CI)	AUC	Sensitivity (95% CI)	AUC	Sensitivity (95% CI)	AUC	Sensitivity (95% CI)
PLCO	CA125	–	0.96	0.86 (0.76, 0.97)	0.75	0.33 (0.13, 0.53)	0.72	0.12 (0.00, 0.27)	0.54	0.03 (0.00, 0.09)
	CA125	0.11	0.94	0.80 (0.68, 0.92)	0.50	0.32 (0.12, 0.51)	0.50	0.12 (0.00, 0.27)	0.50	0.12 (0.01, 0.23)
	HE4	0.19	0.81	0.60 (0.46, 0.74)	0.66	0.23 (0.05, 0.40)	0.55	0.06 (0.00, 0.17)	0.60	0.09 (0.00, 0.18)
	MMP7	0.09	0.65	0.20 (0.08, 0.32)	0.50	0.14 (0.00, 0.28)	0.62	0.18 (0.00, 0.36)	0.48	0.12 (0.01, 0.23)
FHCRC	Mesothelin	0.13	0.72	0.40 (0.26, 0.54)	0.57	0.00	0.62	0.06 (0.00, 0.17)	0.44	0.00
	Spondin-2	0.18	0.61	0.11 (0.02, 0.20)	0.54	0.14 (0.00, 0.28)	0.58	0.06 (0.00, 0.17)	0.57	0.12 (0.01, 0.23)
	SLPI	0.08	0.63	0.16 (0.05, 0.26)	0.53	0.09 (0.00, 0.21)	0.56	0.06 (0.00, 0.17)	0.54	0.03 (0.00, 0.09)
	IGFBP2	0.08	0.64	0.09 (0.01, 0.17)	0.60	0.00	0.49	0.00	0.56	0.03 (0.00, 0.09)
MDACC	B2M	0.17	0.56	0.09 (0.01, 0.17)	0.63	0.05 (0.00, 0.13)	0.60	0.00	0.54	0.15 (0.03, 0.27)
	Transferrin	0.12	0.54	0.09 (0.01, 0.17)	0.53	0.09 (0.00, 0.21)	0.58	0.00	0.53	0.00
	CTAP-III	0.14	0.57	0.09 (0.01, 0.17)	0.56	0.05 (0.00, 0.13)	0.51	0.12 (0.00, 0.27)	0.50	0.09 (0.00, 0.18)
	Hepcidin	0.15	0.53	0.09 (0.01, 0.17)	0.58	0.05 (0.00, 0.13)	0.54	0.12 (0.00, 0.27)	0.51	0.00
Partners	Transferrin	0.12	0.57	0.02 (0.00, 0.07)	0.55	0.05 (0.00, 0.13)	0.55	0.12 (0.00, 0.27)	0.50	0.09 (0.00, 0.18)
	ITIH4	0.26	0.51	0.04 (0.00, 0.10)	0.51	0.05 (0.00, 0.13)	0.59	0.00	0.52	0.03 (0.00, 0.09)
	Apolipoprotein A1	0.17	0.49	0.07 (0.00, 0.14)	0.55	0.05 (0.00, 0.13)	0.51	0.12 (0.00, 0.27)	0.54	0.00
	CA125	0.04	0.96	0.86 (0.76, 0.97)	0.77	0.36 (0.16, 0.56)	0.64	0.06 (0.00, 0.18)	0.53	0.03 (0.00, 0.09)
UPCI	HE4	0.08	0.88	0.73 (0.60, 0.86)	0.65	0.23 (0.05, 0.40)	0.70	0.18 (0.00, 0.36)	0.66	0.12 (0.01, 0.23)
	CA72.4	0.04	0.80	0.44 (0.29, 0.59)	0.59	0.14 (0.00, 0.28)	0.50	0.20 (0.00, 0.40)	0.54	0.06 (0.00, 0.14)
	B7-H4	0.25	0.70	0.36 (0.22, 0.50)	0.56	0.05 (0.00, 0.13)	0.50	0.00	0.47	0.09 (0.00, 0.18)
	CA15.3	0.02	0.72	0.45 (0.31, 0.60)	0.54	0.05 (0.00, 0.13)	0.65	0.27 (0.04, 0.49)	0.52	0.03 (0.00, 0.09)
YUSM	KLK6	0.14	0.73	0.32 (0.18, 0.46)	0.57	0.05 (0.00, 0.13)	0.59	0.07 (0.00, 0.19)	0.57	0.06 (0.00, 0.14)
	CA19.9	0.02	0.50	0.14 (0.03, 0.24)	0.49	0.00	0.66	0.19 (0.00, 0.38)	0.52	0.09 (0.00, 0.18)
	CA125	0.16	0.83	0.27 (0.14, 0.40)	0.39	0.00	0.56	0.00	0.50	0.06 (0.00, 0.14)
	HE4	0.18	0.78	0.31 (0.18, 0.45)	0.47	0.09 (0.00, 0.21)	0.54	0.00	0.57	0.09 (0.00, 0.18)
YUSM	Eotaxin	0.08	0.56	0.02 (0.00, 0.07)	0.60	0.09 (0.00, 0.21)	0.53	0.06 (0.00, 0.17)	0.51	0.09 (0.00, 0.18)
	MMP3	0.09	0.56	0.02 (0.00, 0.07)	0.53	0.05 (0.00, 0.14)	0.50	0.00	0.50	0.15 (0.03, 0.27)
	CA72.4	0.13	0.44	0.02 (0.00, 0.07)	0.46	0.00	0.63	0.00	0.51	0.09 (0.00, 0.18)
	sv-CAM	0.11	0.57	0.04 (0.00, 0.10)	0.54	0.00	0.57	0.12 (0.00, 0.27)	0.50	0.15 (0.03, 0.27)
YUSM	EGFR	0.06	0.52	0.04 (0.00, 0.10)	0.54	0.09 (0.00, 0.21)	0.59	0.06 (0.00, 0.17)	0.52	0.06 (0.00, 0.14)
	Prolactin	0.12	0.54	0.13 (0.03, 0.23)	0.66	0.09 (0.00, 0.21)	0.48	0.12 (0.00, 0.27)	0.50	0.03 (0.00, 0.09)
	CA125	0.29	0.96	0.84 (0.74, 0.95)	0.72	0.23 (0.05, 0.40)	0.58	0.12 (0.00, 0.27)	0.54	0.06 (0.00, 0.14)
	MIF	0.58	0.65	0.18 (0.07, 0.29)	0.48	0.09 (0.00, 0.21)	0.56	0.00	0.57	0.18 (0.05, 0.30)
YUSM	Leptin	0.09	0.52	0.04 (0.00, 0.10)	0.51	0.05 (0.00, 0.13)	0.53	0.12 (0.00, 0.27)	0.54	0.06 (0.00, 0.14)
	Prolactin	0.13	0.55	0.09 (0.01, 0.17)	0.68	0.09 (0.00, 0.21)	0.50	0.06 (0.00, 0.17)	0.56	0.06 (0.00, 0.14)
	OPN	0.54	0.59	0.11 (0.02, 0.20)	0.56	0.05 (0.00, 0.13)	0.55	0.06 (0.00, 0.17)	0.54	0.06 (0.00, 0.14)
	IGF2	0.31	0.60	0.04 (0.00, 0.10)	0.46	0.05 (0.00, 0.13)	0.46	0.06 (0.00, 0.17)	0.49	0.03 (0.00, 0.09)

^aCV were calculated for each of the 60 QC duplicate pairs inserted in the allocated samples and were then averaged over all pairs to obtain the average CV for the assay.

Table 5. Comparing sensitivity at 95% specificity for phase II and III results

Marker	Phase II	Phase III			
		All cases	Cases diagnosed ≤6 mo after draw	Cases diagnosed >6–12 mo after draw	Cases diagnosed >12–18 mo after draw
CA125	0.73	0.86	0.33	0.12	0.03
HE4	0.57	0.73	0.23	0.18	0.12
Transthyretin	0.47	0.02	0.05	0.12	0.09
CA15.3	0.46	0.45	0.05	0.27	0.03
CA72.4 ^a	0.40	0.44	0.14	0.20	0.06
IGFBP2	0.38	0.09	0.00	0.00	0.03
Mesothelin	0.35	0.40	0.00	0.06	0.00
Prolactin	0.34	0.13	0.09	0.12	0.03
Apolipoprotein	0.34	0.07	0.05	0.12	0.19
Spondin-2	0.28	0.11	0.14	0.06	0.12
Transferrin	0.23	0.09	0.09	0.00	0.00
MIF	0.15	0.18	0.09	0.00	0.18
B2M	0.05	0.09	0.05	0.00	0.15

^aAssay changed between phase II and phase III (see Materials and Methods).

markers declined in phase III specimens the further from clinical diagnosis that the specimen was collected.

Discussion

In this study, we assessed how well ovarian cancer biomarkers performed in prediagnostic specimens in asymptomatic women compared with performance in specimens obtained at diagnosis in a different set of subjects. Use of a common set of specimens allowed "head-to-head" comparisons. Generally, markers whose assays had poor CVs also had poor performance as biomarkers. None of the markers with a CV $\geq 30\%$ had a sensitivity (at 95% specificity) better than 37% in either phase II or the phase III data (within 6 months of diagnosis). For "standard" tumor markers with better CVs, such as CA125, HE4, CA72.4, and CA15.3, the performance in phase III data was actually quite comparable to the performance in phase II data when the blood in the cases had been drawn within 6 months of diagnosis. These same markers were observed to have good stability in the paired specimens drawn about a year apart in healthy subjects (see Table 1). In contrast, for markers such as prolactin, transthyretin, or apolipoprotein A1, which may be derived from the patients' response to the cancer, performance was poorer in the phase III specimens even when they were taken within 6 months of the diagnosis compared with that at the time of clinical diagnosis. These were also markers which tended to show less stability in normal paired specimens.

Recently, a panel of markers has been approved to assess cancer risk in women who have an ovarian mass (6). This panel includes several acute phase reactants (apolipoprotein A1, transthyretin, and transferrin) evaluated as part of this study. Although these are analyzed by immunoassays

in the approved panel and by mass spectrometry in this study, inclusion of acute phase reactants raises concern about use of the test to screen for preclinical disease. Physicians and patients should not seek to extend the indication to early detection until it is tested and approved for this indication.

In comparing phases II and III results, differences in the sample sets should be acknowledged, including older age of the phase III subjects, ethnic differences, and more early-stage cases in phase II specimens. We had oversampled early-stage cases, as did the Yale-GOG phase II study (15), with the belief they might provide clues about better markers for detection of preclinical disease. Some of the phase II markers which ranked higher in early-stage cases than in all cases included CA19.9, apolipoprotein A1, and prolactin. However, these markers did not perform well in the phase III specimens, challenging the assumption that markers for early-stage disease are good screening markers. Early-stage ovarian cancer tends to be borderline malignant or low-grade tumors and types such as mucinous that are slower growing and generally have a better prognosis than high-grade serous tumors, relatively few of which are diagnosed at stage I or II.

An important limitation of this study in making inferences about performance of markers relative to CA125 is that CA125 was used in "real time" for triaging women for diagnostic workup for ovarian cancer. Therefore, women with a CA125 value more than 35 may have been selectively excluded at testing points when their CA125 values reached that threshold. In Table 4, it should be noted that the CA125 value used as the cutoff for 95% specificity was 24, not the standard cutoff of 35. When 35 was used as the cutoff, the specificity increased to 98.5% and the sensitivities for time periods 0–6, 6–12, 12–18, and >18 months

decreased to 84%, 19%, 0%, and 3%, respectively. The sensitivities for the more than 6-month intervals, regardless of the cutoff, may be biased downwards because of intervention based on the CA125 level. Also, readers are cautioned not to overinterpret data from Table 4 that the limitations of CA125 (or other markers) can be overcome by screening at more frequent intervals.

Because of this potential bias, phase III specimens which were not used in the setting of clinical decision making are of interest for comparison. A recent study by Anderson and colleagues (22) using specimens from the Carotene and Retinol Efficacy Trial found a similar, though possibly less steep, drop-off in CA125 performance. The AUC for CA125 was 0.74 for cases within 2 years of diagnosis and 0.57 for 4 or more years from diagnosis. Anderson and colleagues concluded that a panel including CA125, HE4, and mesothelin may provide signal for ovarian cancer 3 years before diagnosis and that incorporating prior marker values into the algorithm may also have value (23). Markers showing high stability in healthy controls over time are likely to be the best candidates for this approach which, as noted, tend to exclude the acute phase markers. The longitudinal approach is currently being studied by PLCO and site investigators, and will likely be the subject of future communications.

Relevant to the discussion, the PLCO reported results after 4 rounds of screening with CA125 plus TVU. This regimen produced a high ratio of surgeries to detected cancers (19.5:1) without a clear shift toward earlier stage disease. The authors concluded that screening for ovarian cancer in the general population could not be recommended (23), advice reiterated in several editorials (24, 25). In contrast, Menon and colleagues (26) reported results of a prevalence screen in which two screening modalities were compared: CA125 used as the primary screen with referral for ultrasound if necessary versus ultrasound alone. The trajectory of serial CA125 values was taken into consideration in interpreting positive or negative screening results (25). The sensitivity, specificity, and positive predictive value for ovarian and tubal cancers were all higher with CA125 followed by ultrasound (89.4%, 99.8%, and 43.3%) than with ultrasound alone (75.0%, 98.2%, and 5.3%). The ratio of operations per malignancy found within one year of the prevalence screen was 2.3 for CA125 but 18.8 for TVU, the latter value being similar to what was observed in the PLCO. The authors concluded that general population screening using CA125 was "feasible" (26).

In summary, we tested 28 ovarian cancer biomarkers in prediagnostic specimens from the PLCO. CA125

remains the single-best biomarker for ovarian cancer and has its strongest signal within 6 months of diagnosis. Though disappointing, this conclusion should be viewed in perspective of lessons learned from the study and future directions suggested. Refinement of assays with poor CV is desirable prior to phase III testing of a biomarker because, without this, its performance may not be fairly tested. Performance of a biomarker in phase II provides clues about performance in phase III when the phase III specimen was taken within 6 months of diagnosis and when the marker does not represent an acute reaction to clinical disease. Whether the decline in performance of the current best ovarian cancer biomarkers in specimens more than 6 months remote from diagnosis is a limitation of screening studies, in which CA125 prompted clinical action or represents an inherent limitation due to a short lead time for ovarian cancer, requires further study. If there is hope for reduced mortality for ovarian cancer through screening, we need markers that would show a signal for ovarian cancer more than 6 months remote from diagnosis. It may be especially worthwhile to focus discovery efforts on high-grade invasive tumors associated with a normal CA125. Even phase II specimens with these characteristics will be valuable because cases with low CA125 at clinical diagnosis would likely have had a low CA125 during their presymptomatic phase.

Disclosure of Potential Conflicts of Interest

R.C. Bast reports royalties for CA125 and is a member of the Scientific Advisory Boards for Vermillion and for Fujirebio Diagnostics Inc.

Acknowledgments

The authors thank Allison Vitonis for programming support.

Grant Support

This work was supported in part by the Ovarian Cancer SPOREs at FHCRC (P50 CA083636), to N. Urban, MDACC (P50 CA83639) to R.C. Bast and FCCC/UPenn (P50 CA083638), an NCI contract (N01-CN-43309), an EDRN CVC (5U01 CA113916) to A.K. Godwin; an Ovarian Cancer SPORE at BWH (P50 CA105009) and an EDRN (5U01 CA86381) to D.W. Cramer; and grants from Golfers Against Cancer and the Wiley Mossy Foundation to R.C. Bast.

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 August 10, 2010; revised November 22, 2010; accepted December 8, 2010; published OnlineFirst February 24, 2011.

References

1. Terry KL, Sluss PM, Skates SJ, Mok SC, Ye B, Vitonis AF, et al. Blood and urine markers for ovarian cancer: a comprehensive review. *Dis Markers* 2004;20:53-70.
2. Gorelik E, Landsittel DP, Marrangoni AM, Modugno F, Velikokhatnaya L, Winans MT, et al. Multiplexed immunobead-based cytokine profiling for early detection of ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 2005;14:981-7.
3. Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, et al. Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 2002;359:572-7.

4. Visintin I, Feng Z, Longton G, Ward DC, Alvero AB, Lai Y, et al. Diagnostic markers for early detection of ovarian cancer. *Clin Cancer Res* 2008;14:1065–72.
5. Zhang Z, Bast RC Jr., Yu Y, Li J, Sokoll LJ, Rai AJ, et al. Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res* 2004;64:5882–90.
6. Fung ET. A recipe for proteomics diagnostic test development: the OVA1 test, from biomarker discovery to FDA clearance. *Clin Chem* 2010;56:327–9.
7. Jacobs IJ, Skates SJ, MacDonald N, Menon U, Rosenthal AN, Davies AP, et al. Screening for ovarian cancer: a pilot randomised controlled trial. *Lancet* 1999;353:1207–10.
8. Gohagan JK, Prorok PC, Hayes RB, Kramer BSP, Prostate LCaOCSTPT. The Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial of the National Cancer Institute: history, organization, and status. *Control Clin Trials* 2000;21:251S–72S.
9. Hasson MA, Fagerstrom RM, Kahane DC, Walsh JH, Myers MH, Caughman C, et al. Design and evolution of the data management systems in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials* 2000;21:329S–48S.
10. Hayes RB, Reding D, Kopp W, Subar AF, Bhat N, Rothman N, et al. Etiologic and early marker studies in the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial. *Control Clin Trials* 2000;21:349S–55S.
11. Miller AB, Yurgalevitch S, Weissfeld JL, Prostate LCaOCSTPT. Death review process in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials* 2000;21:400S–6S.
12. O'Brien B, Nichaman L, Browne JE, Levin DL, Prorok PC, Gohagan JK, et al. Coordination and management of a large multicenter screening trial: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials* 2000;21:310S–28S.
13. Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, et al. Phases of biomarker development for early detection of cancer. *J Natl Cancer Inst* 2001;93:1054–61.
14. Mor G, Visintin I, Lai Y, Zhao H, Schwartz P, Rutherford T, et al. Serum protein markers for early detection of ovarian cancer. *Proc Natl Acad Sci U S A* 2005;102:7677–82.
15. Mor G, Symanowski J, Visintin I, Birrer MJ, Ward D. Six Biomarker Panel detects early stage ovarian cancer. 2009AACR Annual Meeting Denver, CO; 2009.
16. Scholler N, Garvik B, Quarles T, Jiang S, Urban N. Method for generation of in vivo biotinylated recombinant antibodies by yeast mating. *J Immunol Methods* 2006;317:132–43.
17. Scholler N, Lowe KA, Bergan LA, Kampani AV, Ng V, Forrest RM, et al. Use of yeast-secreted in vivo biotinylated recombinant antibodies (Biobodies) in bead-based ELISA. *Clin Cancer Res* 2008;14:2647–55.
18. Diamandis EP, Scorilas A, Fracchioli S, Van Gramberen M, De Bruijn H, Henrik A, et al. Human kallikrein 6 (hK6): a new potential serum biomarker for diagnosis and prognosis of ovarian carcinoma. *J Clin Oncol* 2003;21:1035–43.
19. Nolen B, Marrangoni A, Velikokhatnaya L, Prosser D, Winans M, Gorelik E, et al. A serum based analysis of ovarian epithelial tumorigenesis. *Gynecol Oncol* 2009;112:47–54.
20. Pepe MS, Longton G. Standardizing diagnostic markers to evaluate and compare their performance. *Epidemiology* 2005;16:598–603.
21. Zhu CS, Pinsky PF, Cramer DW, et al. A framework for evaluating biomarkers for early detection: validation of biomarker panels for ovarian cancer. *Can Prev Res* 2011;4:375–383.
22. Anderson GL, McIntosh M, Wu L, Barnett M, Goodman G, Thorpe JD, et al. Assessing lead time of selected ovarian cancer biomarkers: a nested case-control study. *J Natl Cancer Inst* 2010;102:26–38.
23. Partridge E, Kreimer AR, Greenlee RT, Williams C, Xu JL, Church TR, et al. Results from four rounds of ovarian cancer screening in a randomized trial. *Obstet Gynecol* 2009;113:775–82.
24. Clarke-Pearson DL. Clinical practice. Screening for ovarian cancer. *N Engl J Med* 2009;361:170–7.
25. Mutch DG. Ovarian cancer: to screen or not to screen. *Obstet Gynecol* 2009;113:772–4.
26. Menon U, Gentry-Maharaj A, Hallett R, Ryan A, Burnell M, Sharma A, et al. Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *Lancet Oncol* 2009;10:327–40.