

NIH Public Access

Author Manuscript

Cancer Prev Res (Phila). Author manuscript; available in PMC 2012 March 1.

Published in final edited form as:

Cancer Prev Res (Phila). 2011 March ; 4(3): 365–374. doi:10.1158/1940-6207.CAPR-10-0195.

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

Daniel W. Cramer1, **Robert C. Bast Jr.**2, **Christine D. Berg**3, **Eleftherios P. Diamandis**4, **Andrew K. Godwin**5, **Patricia Hartge**6, **Anna E. Lokshin**7, **Karen H. Lu**2, **Martin W. McIntosh**8, **Gil Mor**9, **Christos Patriotis**10, **Paul F. Pinsky**3, **Mark D. Thornquist**8, **Nathalie Scholler**11, **Steven J. Skates**12, **Patrick M. Sluss**12, **Sudhir Srivastava**10, **David C. Ward**13, **Zhen Zhang**14, **Claire S. Zhu**3, and **Nicole Urban**⁸

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 ¹²Massaschusetts General Hospital, Boston, Massachusetts ¹³Cancer Research Center, University of Honolulu, Hawaii ¹⁴Johns Hopkins University, Baltimore, Maryland

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 (EDRN) or Ovarian Cancer Specialized Program of Research Excellence (SPORE) 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 six 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.

Correspondence to: Daniel W. Cramer, MD, Brigham and Women's Hospital, Ob/Gyn, Division of Epidemiology, 221 Longwood Ave., RFB 365, Boston, MA 02115, dcramer@partners.org, Tel: 617-732-4895, Fax: 617-732-4899.

Conflicts of Interest RCB reports royalties for CA125 and is a member of the Scientific Advisory Boards for Vermillion and for Fujirebio Diagnostics Inc.

Keywords

Ovarian neoplasms; CA125; HE4; Screening tests; CA72.4

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 CA125 and 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 pre-operative 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 four 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 five identical specimen sets to investigators based upon 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 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 Hutchison 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). About half 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, Norton, MA) were included. Cases had blood drawn prior to surgery or chemotherapy, processed generally within four 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 ml aliquots of sera for each subject were assembled and sent to FCCC (AKG) for aliquot construction and re-labeling 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 performed by the EDRN Coordinating Center (MDT).

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 (14), modified to include CA125, and developed into a kit (Cancer Panel) produced by Millipore. This five marker panel was evaluated in 156 newlydiagnosed 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 stage I, II and 100 stage III, 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 of 2006, 118 cases of invasive ovarian, primary peritoneal, and fallopian tube cancers with available samples had been confirmed. Tumors of borderline malignancy were excluded. For cases a single serum sample most proximate to diagnosis was selected. Controls were selected from healthy individuals who remained cancer free and matched to cases by five-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 0.6ml, 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 five separate laboratories by individuals blind to case or control status. Investigators at the FHCRC (NU, NS, MWM) evaluated 5 markers using single-plex Luminex bead assays, as described for CA125 (16) HE4 and mesothelin (17). Reagents supplied by Diadexus were used for Spondin-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 MMP-7 (R&D Systems kit) were introduced. The SLPI assay was performed using a pool of recombinant antibodies for capture and goat anti-SLPI pAb (R&D Systems) for detection. Investigators at MDACC (RCB, KHL) and Johns Hopkins University (JHU: ZZ) evaluated 7 markers using a mass spectrometer-based system for its

markers in both phases further described in (5). Investigators at Partners (DWC, PMS, SJS) and at Mount Sinai (Ontario: EPD) evaluated 13 markers using plate-based assay for B7-H4 (Diadexus), DcR3 (Diadexus), Macrophage inhibitory substance 1 (MIC1) (DiaDexus), Spondin-2 (DiaDexus), IGF II (Diagnostic Systems Laboratories, Beckman), Mesothelin (SMRP: Fujirebio), HE4 (Fujirebio), CH1311 (YKL-40, Chitinase; Quidel), and Kallikrein 6 (as described in (18)). CA72.4 was performed by IRMA (Fujirebio). Platform-based assays (Modular E170, Roche) were used for CA125, CA15.3, CA19.9, and 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 UPCI (AL) evaluated 34 markers in phase II using a multi-plex bead assay system described in (19). For phase III, the number of markers was reduced to 8 but the same technique used. These included: CA125, HE4, Eotoxin, MMP3, CA72.4, soluble Vascular Cell Adhesion Molecule 1 (sV-Cam), Epidermal Growth Factor Receptor (EGFR), and Prolactin. Investigators at YUSM (GM, DCW) evaluated 6 markers in their phase II study and phase III study using a multi-plex bead system described in (4). These markers included CA125, Macrophage inhibiting factor, 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 EDRN. 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. Based on the common phase II results, the PLCO provided the phase III proximate samples in a blinded manner to the assay sites (FHCRC, MDA, 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 (PFP). 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 (CVs) and correlations based upon the paired specimens. In general, CVs were greatest for multi-plex Luminex assays, less for single-plex Luminex or plate-based assays, and least for platform-based assays of the type used in clinical laboratories. The most stable markers, based upon the 40 serial paired specimens were CA125, AFP, CA15.3, CA19.9, CA72.4, mesothelin, and HE4. The least stable markers were prolactin, cytokeratin 19, growth hormone, myeloperoxidase, tumor necrosis factor receptor 2, and kallikrein 6.

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 post-menopausal) 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 twenty 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 out of the 49 markers originally tested. The markers not shown all had sensitivities $\langle 0.17 \rangle$ and AUCs $\langle 0.64 \rangle$ 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 pre-menopausal women or cases with borderline malignancies did not substantially change the all case rankings.

Turning to phase III data, results are shown for the five 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 re-measured CA125) performed best. The sensitivity (and 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 (and 95% CI), among cases ≤ 6 months from diagnosis, was 0.73 (0.60-0.86). Like CA125, its sensitivity declined in specimens more remote than six 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, like 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 like prolactin, transthyretin, and transferrin. The performance of nearly all 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 pre-diagnostic 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-tohead" comparisons. Generally, markers whose assays had poor CVs also had poor performance as biomarkers. None of the markers with a $CV \ge 30\%$ 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, like 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 like 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 to 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 pre-clinical 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 phase 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-staged 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-staged 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 like 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 of >35 may have been selectively excluded at testing points when their CA125 values reached that threshold. It should be noted in Table 4 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 over-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 over-interpret 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 et al (22) using specimens from the Carotene and Retinol Efficacy Trial (CARET) found a similar, though possibly less steep, drop-off in CA125 performance. The AUC for CA125 was 0.74 for cases within two years of diagnosis and 0.57 for 4 or more years from diagnosis. Anderson et al concluded that a panel including CA125, HE4, and mesothelin may provide signal for ovarian cancer three 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 four rounds of screening with CA125 plus transvaginal ultrasound. This regimen produced a high ratio of surgeries to detected cancers (19.5 to 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 et al. 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 vs. 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 transvaginal ultrasound, the latter value being similar to what was observed in the PLCO. The authors concluded general population screening using CA125 was "feasible" (26).

In summary, we tested 28 ovarian cancer biomarkers in pre-diagnostic 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 desireable prior to phase III testing of a biomarker since, without this, its performance may not be fairly tested. Performance of a biomarker in phase II does provide 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 where 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 since cases with low CA125 at clinical diagnosis would likely have had a low CA125 during their pre-symptomatic phase.

Acknowledgments

Funding This work was supported in part by the Ovarian Cancer SPOREs at FHCRC (P50 CA083636), to NU, MDACC (P50 CA83639) to RCB and FCCC/UPenn (P50 CA083638), an NCI contract (N01-CN-43309), an EDRN CVC (5U01 CA113916) to A.K.G.; an Ovarian Cancer SPORE at BWH (P50 CA105009) and an EDRN (5U01 CA86381) to D.W.C., and grants from Golfers Against Cancer and the Wiley Mossy Foundation to RCB. Programming support by Allison Vitonis is gratefully acknowledged.

References

- 1. Terry KL, Sluss PM, Skates SJ, et al. Blood and urine markers for ovarian cancer: a comprehensive review. Dis Markers. 2004; 20:53–70. [PubMed: 15322314]
- 2. Gorelik E, Landsittel DP, Marrangoni AM, et al. Multiplexed immunobead-based cytokine profiling for early detection of ovarian cancer. Cancer Epidemiol Biomarkers Prev. 2005; 14:981–7. [PubMed: 15824174]
- 3. Petricoin EF, Ardekani AM, Hitt BA, et al. Use of proteomic patterns in serum to identify ovarian cancer. Lancet. 2002; 359:572–7. [PubMed: 11867112]
- 4. Visintin I, Feng Z, Longton G, et al. Diagnostic markers for early detection of ovarian cancer. Clin Cancer Res. 2008; 14:1065–72. [PubMed: 18258665]
- 5. Zhang Z, Bast RC Jr, Yu Y, et al. Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. Cancer Res. 2004; 64:5882–90. [PubMed: 15313933]
- 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. [PubMed: 20110452]
- 7. Jacobs IJ, Skates SJ, MacDonald N, et al. Screening for ovarian cancer: a pilot randomised controlled trial. Lancet. 1999; 353:1207–10. [PubMed: 10217079]
- 8. Gohagan JK, Prorok PC, Hayes RB, Kramer BS. 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. [PubMed: 11189683]
- 9. Hasson MA, Fagerstrom RM, Kahane DC, 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. [PubMed: 11189686]
- 10. Hayes RB, Reding D, Kopp W, 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. [PubMed: 11189687]

- 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. [PubMed: 11189691]
- 12. O'Brien B, Nichaman L, Browne JE, 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. [PubMed: 11189685]
- 13. Pepe MS, Etzioni R, Feng Z, et al. Phases of biomarker development for early detection of cancer. J Natl Cancer Inst. 2001; 93:1054–61. [PubMed: 11459866]
- 14. Mor G, Visintin I, Lai Y, et al. Serum protein markers for early detection of ovarian cancer. Proc Natl Acad Sci U S A. 2005; 102:7677–82. [PubMed: 15890779]
- 15. Mor, G.; Symanowski, J.; Visintin, I.; Birrer, MJ.; Ward, D. Six Biomarker Panel detects Early Stage Ovarian Cancer. 2009 AACR 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. [PubMed: 17113097]
- 17. Scholler N, Lowe KA, Bergan LA, et al. Use of yeast-secreted in vivo biotinylated recombinant antibodies (Biobodies) in bead-based ELISA. Clin Cancer Res. 2008; 14:2647–55. [PubMed: 18451228]
- 18. Diamandis EP, Scorilas A, Fracchioli S, 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. [PubMed: 12637468]
- 19. Nolen B, Marrangoni A, Velikokhatnaya L, et al. A serum based analysis of ovarian epithelial tumorigenesis. Gynecol Oncol. 2009; 112:47–54. [PubMed: 19007974]
- 20. Pepe MS, Longton G. Standardizing diagnostic markers to evaluate and compare their performance. Epidemiology. 2005; 16:598–603. [PubMed: 16135934]
- 21. Zhu CS, Pinsky PF, Cramer DW, et al. A Framework for Evaluating Biomarkers for Early Detection: Validation of Biomarker Panels for Ovarian Cancer. submitted.
- 22. Anderson GL, McIntosh M, Wu L, et al. Assessing lead time of selected ovarian cancer biomarkers: a nested case-control study. J Natl Cancer Inst. 2010; 102:26–38. [PubMed: 20042715]
- 23. Partridge E, Kreimer AR, Greenlee RT, et al. Results from four rounds of ovarian cancer screening in a randomized trial. Obstet Gynecol. 2009; 113:775–82. [PubMed: 19305319]
- 24. Clarke-Pearson DL. Clinical practice. Screening for ovarian cancer. N Engl J Med. 2009; 361:170– 7. [PubMed: 19587342]
- 25. Mutch DG. Ovarian cancer: to screen or not to screen. Obstet Gynecol. 2009; 113:772–4. [PubMed: 19305318]
- 26. Menon U, Gentry-Maharaj A, Hallett R, 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. [PubMed: 19282241]

Table 1

Coefficients of variation based on 84 replicates and Pearson correlations (in parentheses) based on 40 paired serial samples for assays in the common phase II set by site.***

*** Partners performed platform (Roche E170) or plate-based assays; FHCRC performed single-plex Luminex bead assays and plate-based assays; MDACC performed a mass spectographic analysis using surface-enhanced antibody chips; UPCI performed multiplexed Luminex bead assays.

NIH-PA Author Manuscript

Table 2

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

 NIH-PA Author ManuscriptNIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Phase II results for all cases and early stage cases compared to general population controls by sensitivity at 95% specificity and area under the curve Phase II results for all cases and early stage cases compared to general population controls by sensitivity at 95% specificity and area under the curve (sorted by sensitivity for all cases). (sorted by sensitivity for all cases).

94 (SS:0-05:0) IS:0 37 (0.1:0-00:0) 19:0 6.6 (2:0-0-IS:0)LS:0 LS: (9.1:0-50:0) 60:0 B2M $B2M$ $0.50 (0.03 (0.02 - 0.12)$ 45 $0.56 (0.51 - 0.61)$ 41 $0.05 (0.00 - 0.12)$ 47 $0.50 (0.50 - 0.58)$ 49

NIH-PA Author Manuscript

 $\ddot{6}$

 $0.50(0.50-0.58)$

 $0.05(0.00 - 0.12)$

 $\overline{4}$

 $0.56(0.51 - 0.61)$

 45

 $0.05\ (0.02\text{--}0.12)$

 $\frac{4}{6}$ 27

> 48 \ddot{t}

 NIH-PA Author ManuscriptNIH-PA Author Manuscript

Coefficients of variation (CV) 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.

 NIH-PA Author ManuscriptNIH-PA Author Manuscript

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

