Published in final edited form as:

Int J Gynecol Cancer. 2016 July ; 26(6): 1070–1077. doi:10.1097/IGC.0000000000000737.

Validation of a biomarker panel and longitudinal biomarker performance for early detection of ovarian cancer

Archana R. Simmons, PhD¹, Charlotte H. Clarke, PhD¹, Donna B. Badgwell, PhD¹, Zhen Lu, M.D.¹, Lori J. Sokoll, PhD², Karen H. Lu, M.D.³, Zhen Zhang, PhD^{2,*}, Robert C. Bast Jr., M.D.^{1,*}, and Steven J. Skates, PhD^{4,*}

¹Department of Experimental Therapeutics, University of Texas M.D. Anderson Cancer Center, Houston, TX 77030

²Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD, 21287

³Department of Gynecologic Oncology, University of Texas M.D. Anderson Cancer Center, Houston, TX 77030

⁴Department of Biostatistics, Massachusetts General Hospital, Boston, MA 02114

Abstract

Objectives—Longitudinal multi-marker combinations have the potential to improve sensitivity while maintaining the high specificity required for early detection of ovarian cancer. The use of multiple markers to improve sensitivity over CA125 in longitudinal algorithms for early ovarian cancer detection requires the selection of markers with optimal discriminatory power and low longitudinal variance relative to disease-initiated changes. Our objective was to identify a multi-marker panel suitable for ovarian cancer, where each individual marker has its own baseline, permitting longitudinal algorithm development.

Materials and methods—In this retrospective study, we measured CA125, HE4, MMP-7, CA72-4, CA19-9, CA15-3, CEA and s-VCAM concentrations using immunoassays in pre-treatment sera from 142 stage I ovarian cancer cases and 5 annual samples each from 217 healthy controls. Following random division into training and validation sets, all possible biomarker combinations were exhaustively explored using linear classifiers, to identify the panel with greatest sensitivity for stage I disease at a high specificity of 98%. To evaluate longitudinal performance of the individual markers, the within-person over time and the between-person coefficient of variation (CV) were estimated. Hierarchical modeling across women of log-concentrations enabled borrowing of information across subjects to moderate variance estimates, given the small number of observations per subject.

Results—The four marker panel comprising CA125, HE4, MMP-7 and CA72-4 performed with the highest sensitivity (83.2%) at 98% specificity. The within-person CVs were lower for

Address correspondence to: Robert C. Bast Jr., M.D., UT MD Anderson Cancer Center, 1400 Pressler Street, Unit 1439, Houston, TX 77230-1439; Phone 713-792-7743; Fax 713-792-7864; rbast@mdanderson.org.

*These authors contributed equally to the manuscript.

Disclosures and Potential Conflicts of Interest

Dr. Bast receives royalties from Fujirebio Diagnostics, Inc. for discovery of CA125 and is a member of the advisory board for Vermillion.

CA125, HE4, MMP-7 and CA72-4 (15%, 25%, 25% and 21% respectively) compared to their corresponding between-person CV (49%, 20%, 35% and 84% respectively) indicating baselines in healthy volunteers. Following simple log-transformations, within-volunteer variation across volunteers was modeled with a normal distribution, permitting parsimonious hierarchical modeling.

Conclusions—The multiplex panel chosen is suitable for early detection of ovarian cancer and the individual markers have their own baseline permitting longitudinal algorithm development.

INTRODUCTION

In 2016, an estimated 22,280 women in the United States will be diagnosed with ovarian cancer and 14,240 women will die from this disease [1]. For early stage cancers, survival rates range between 73–92%, whereas in advanced stage cancers, long-term survival rates are less than 30% [2]. Only a small percentage of women (~32%) are diagnosed at an early stage, whereas most (~62%) are diagnosed at an advanced stage, when the disease is no longer confined to the ovary or pelvis [2]. Early detection through screening has the potential to improve survival. However, no test is recommended for general population screening. The low prevalence of ovarian cancer (1 in 2500 post-menopausal women) requires stringent diagnostic performance for any screening strategy to attain a minimum acceptable positive predictive value (PPV) of 10% (99.8% specificity, >75% sensitivity), where at most 10 operations would need to be performed for each detected case [3]. Serum Cancer Antigen 125 (CA125) and Transvaginal Ultrasound (TVS) are recommended in women with suspicious pelvic masses to prompt referral to gynecologic oncologists [3]. However, a strategy where CA125 interpreted at a single time point prompts ultrasound does not achieve adequate sensitivity for early stage detection [4], while simultaneous first line testing with both modalities, interpreted at a single time point does not provide adequate specificity [5].

Serum CA125 levels monitored over time remain stable in healthy individuals, but rise steadily above a woman's individual baseline in most women that develop ovarian cancer [6]. When these serial values were interpreted with the Risk of Ovarian Cancer Algorithm (ROCA), the strategy achieved a sensitivity of 86% at 98% specificity, in contrast with 62% at the same specificity for fixed CA125 cut-offs [6]. ROCA is now being evaluated as a first-line screen prompting TVS as a second-line screen, as a part of a multi-modal screening (MMS) strategy in the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) [7] and in the Normal Risk Ovarian Cancer Screening Study (NROSS) in the United States [8]. With each new CA125 test, a ROCA score is calculated, and women are triaged into a low-risk group (annual CA125 screening referral), intermediate-risk group (3-month repeat CA125 referral) and a high-risk group (referral to TVS and a gynecological oncologist and if necessary, surgery) [7]. Triaging is revised based on the new ROCA score for every new CA125 result. Based on this strategy, the prevalence screen of UKCTOCS achieved a PPV of 43.3%, indicating only 3 operations per case of ovarian cancer detected. A similar PPV, substantially higher than the required minimum of 10%, was also observed in NROSS, indicating strong promise for a ROCA-based MMS strategy [8]. Recent results

suggest promising evidence of mortality reduction using MMS in the UKCTOCS, although further follow-up is needed [9].

Despite the promise of CA125-based MMS, CA125 is expressed in only 80% of ovarian cancers and additional markers will be needed to improve upon sensitivity, while maintaining specificity [10]. A panel comprising four biomarkers has been recommended to achieve greater coverage of the heterogeneity of ovarian cancer [11]. Several biomarker panels have been proposed for improved sensitivity over CA125 alone [11]. We investigated 96 promising serum biomarkers utilizing multiplex xMAP bead-based immunoassays to identify several promising four-marker [12]. One such panel comprising CA125, HE4, CEA and sVCAM-1 yielded a sensitivity of 86% at 98% specificity for early stage cancers [12].

A longitudinal multi-marker risk of ovarian cancer algorithm (MROCA) benefits from enhanced sensitivity of multiple markers to cover a greater spectrum of disease. In addition, the MROCA benefits from serial measurements to improve sensitivity for detection earlier in time. Thus, the MROCA may improve upon the CA125-based ROCA while maintaining the same high specificity. In our previous work, assessment of 96 biomarkers utilizing multiplex bead-based immunoassays led to the identification of several promising candidate biomarker panels. These panels had similar performance for distinguishing early stage ovarian cancer cases from controls with 86% sensitivity at 98% specificity [12]. The simultaneous measurement in the same volume of serum of multiple markers using multiplexed assays, though conducive to high-throughput discovery, led to high assay CVs precluding definitive selection of biomarkers. Our goal in this study was to utilize standard quantitative immunoassays and automated clinical platforms (when available) with substantially lower assay CVs compared to multiplex assays, to assess individual sera for each biomarker, with the ultimate goal of establishing an optimal multi-marker panel for early ovarian cancer detection.

In this current study, we have evaluated the top eight candidates from our previous efforts using standard ELISA assays, which have lower assay CVs than bead-based assays, and identified an optimal linear classifier based on a four marker panel for early detection. We have assessed the within-person variation over time and the between-person variation of these biomarkers in healthy individuals to evaluate their potential as longitudinal biomarkers and to derive initial modeling parameters towards development of MROCA for early detection of ovarian cancer.

MATERIALS AND METHODS

Study Population

Pre-treatment sera from 142 stage I ovarian cancer cases (one per case, ages 20–87) from the Gynecologic Oncology Group Trial (GOG) and five annual serum samples from 217 healthy post-menopausal women (ages 55–85) enrolled in NROSS in whom ovarian cancer was not detected formed cases and controls respectively in the sample set. Sample characteristics are summarized in Table 1.

Sample collection, storage and assay

Serum sample collection and storage followed standard IRB approved protocols outlined in the NROSS study. Sera were stored at -80°C prior to biomarker measurements.

Immunoassays were utilized to measure biomarker concentrations using manufacturer recommended protocols; Roche (CA125, CA19-9, CEA, CA15-3, CA72-4), Fujirebio Diagnostics (HE4), R&D Systems (MMP-7, s-VCAM). The Roche assays were performed on the Elecsys 2010 analyzer and the others were performed in the ELISA format.

Statistical Analysis

A repeated sub-sampling validation procedure was used to identify an optimal panel among all possible linear combinations of the eight markers. The marker concentrations were first log-transformed. During each sub-sampling, the samples were randomly divided into a training set (to model the data) comprising 60% of the samples and a validation set (to test the model) with the remaining 40%. The 60–40 split was chosen to minimize the validation error and the error rate on the training set. For a particular marker panel, a linear classifier was derived using the training data. Panels and associated classifiers that had sensitivity greater than 82% at 98% specificity on training samples were further selected. This was based on the derived classifiers' mean sensitivity at 98% specificity on test samples. This was completed over 20 repeated sub-sampling runs. Initially, in addition to linear models, quadratic classifiers were also evaluated. The resulting improvement in model performance was minimal and therefore not included in the reported results.

The distributions of candidate markers, both within-woman over time, and between-women, were examined using the Box-Cox transformation, to determine a scale on which standard Gaussian models fit best. This approach identifies the optimal power-normalized transformation, with a power of zero corresponding to a log-transformation. The best integer transformation for all candidates was the log-transformation except for CA72.4, where no one power transformation resulted in a Gaussian distribution being a good approximation. An advantage with the log-transformation is the interpretation of the standard deviation (SD) which to first order approximation is the CV on the original scale. Further investigation of transformations for CA72-4 revealed that a $\log(\log(\bullet)+1)$ -transformation gave approximate symmetry, suitable for Gaussian distribution modeling. Following identification of the best normalizing transformation for each biomarker, hierarchical Gaussian models were implemented in WinBUGS [13] to estimate control subjects' individual means and SDs over time, and the SD between the subjects' means.

RESULTS

Identification of Multiplex Biomarker Panel

CA125, HE4, MMP-7, s-VCAM, CA15-3, CA 19-9, CA 72-4 and CEA were evaluated with standard immunoassays in the pre-operative sera of 142 stage I ovarian cancers and five longitudinal serial samples from 217 healthy women (1085 total samples) who did not develop cancer. The immunoassay results obtained are shown in Fig.1 as dot plots. s-VCAM was down-regulated in cases in comparison to controls whereas all other markers were up-regulated in the cases.

Among individual markers, CA125 had the highest sensitivity at high specificities followed by HE4 as expected based on previous studies [14]. The multi-marker panels were chosen to improve sensitivity over that of CA125 at the high specificity of 98%. The top-ranked panels of four or less markers are listed in Table 2 in order of their mean sensitivity in test samples over the repeated sub-sampling runs. Based on the two highest sensitivities (83.7%, 83.2%) at 98% specificity in the validation sets, a panel comprising CA125, CA72-4 and HE4 with the fourth marker as s-VCAM or MMP-7 was chosen to be an optimal panel. The simplicity of linear classifiers helped to ensure the generalizability of model performance in test samples. In addition, given the relatively large sample size of the study, noticeably small variances were obtained for test sample sensitivities estimated through repeated sub-sampling.

Modeling longitudinal biomarker behavior

Though interpreting CA125 at a single time point does not have sufficient sensitivity for detection of most ovarian cancers in early stage, and is therefore not utilized, longitudinal evaluation of CA125 has greater sensitivity at the same specificity and has shown substantial promise as the first line test in a multimodal early detection strategy. This increase in sensitivity is due to the relatively stable levels of CA125 over time in women that are apparently free of ovarian cancer and rise significantly above a woman's baseline over time in women that are subsequently diagnosed with ovarian cancer. The next step was to identify if the candidate biomarkers in the optimal panel provide additional utility when interpreted longitudinally and to estimate the parameters describing their longitudinal behavior. Two separate statistical models are needed to calculate the risk of having ovarian cancer using longitudinal behavior of biomarkers: (i) models for stable longitudinal biomarker levels in apparently healthy women and; (ii) models for longitudinal biomarker levels that increase after a change-point above a woman's baseline in women that subsequently are diagnosed with ovarian cancer. For healthy women apparently free of ovarian cancer, this model comprises; (a) the distribution between-women of the individual baseline level described by the population mean level (between-person CV) (b) the distribution of the variation about the baseline level (within-person CV). Low within-person CV compared to the average between-person CV identifies a candidate marker with promise for longitudinal interpretation to show substantial increases in sensitivity compared to interpreting the candidate marker at a single time point at the same specificity.

Within-person variation

For biomarkers to exhibit longitudinal utility, each participant should have her own baseline for the biomarker. To understand if such baseline levels exist for the markers in this study, we evaluated the within-person over time and the between-person CV for each biomarker on the log-concentration scale utilizing five longitudinal annual serial samples from 217 healthy individuals who did not develop cancer. The results are summarized in Table 3.

All chosen biomarkers have a low median within-person CV in healthy individuals indicative of low within-person biological variability. In other words, for these markers, each apparently healthy individual has their own baseline, a characteristic necessary for a candidate marker to increase sensitivity when interpreted longitudinally. In the chosen panel,

CA125 had the lowest median biological within-person over time variability, followed by CA 72-4 and HE4. While MMP-7 and s-VCAM offer similar power for discriminating cases from controls in a panel, MMP-7 had lower within-person CV (20.2%) in comparison to sVCAM (22.8%); hence MMP-7 was assigned a higher rank. On this basis, the final panel was comprised of CA125, HE4, CA 72-4 and MMP-7. The lower quantiles of the biological within-person variability also serves as an important upper limit for the analytical performance metrics to guide assay development for candidates in the panel. Analytical-variability measured in terms of intra- and inter-assay CVs should be lower than the biological-variability over time within apparently healthy individuals. As interpretation of biological variation within a woman's biomarker levels over time is crucial to the development of longitudinal biomarker algorithms, lower assay variability will not significantly add to existing biological variability and therefore minimize the uncertainty contributed by analytical noise.

The distribution of the within-person SD across volunteers is plotted in Fig.2 on a log-scale. The SDs of CA125, HE4 and MMP-7 follow a simple Gaussian distribution permitting the distribution of the CV across women to be modeled hierarchically with a simple Gaussian distribution, enabling borrowing of information across women. For CA72-4, an additional log-transformation was required for the within-person variation to follow a Gaussian distribution.

Between-person variation

For the biomarkers CA125, HE4, MMP-7 and log CA72-4, the median between-person variation was evaluated utilizing WINBUGS using the longitudinal healthy controls and the data is tabulated in Table 4. All markers with the exception of HE4 had significantly lower within-person variability in comparison to the between-person variability as evidenced by the low within/between ratios, further confirming the existence of a within-woman baseline conducive to the development of longitudinal algorithms.

DISCUSSION

Use of ROCA to evaluate CA125 over time in MMS, has provided greater sensitivity and promising mortality reductions for detecting ovarian cancer, than single time-point CA125 in both the UKCTOCS and the NROSS trials [9]. Nonetheless, at least 20% of ovarian cancers will not be detected at an early stage by CA125 alone. A panel of biomarkers evaluated longitudinally using MROCA has the potential to enhance sensitivity by detecting cases missed by CA125 and improving lead times, thereby enhancing sensitivity for early stage disease, while maintaining the high specificity required to achieve the minimum acceptable PPV.

The goal here was to establish a biomarker panel from a subset of eight promising biomarkers that would improve upon the performance of CA125 alone and to determine whether healthy postmenopausal women would exhibit an individual baseline for each of these biomarkers (where the individual-variation over time would be less than the between-individual variation). Having established the within-person variation over time, between-person variation and the distribution of these variations, we have established the

transformations and the parameter estimates for control subjects in MROCA. These models describing the longitudinal-behavior of multiple-markers in women that do not develop ovarian cancer form a vital component of the statistical modeling in the development of the MROCA.

While all biomarker candidates were either significantly up-regulated or down-regulated in ovarian cancer cases compared to controls, a combination of biomarkers including CA125, HE4, CA72-4 and MMP-7 emerged as the top panel with a sensitivity of 83.2% at 98% specificity in the validation tests, improving on CA125 alone. The CA125 and HE4 combination achieved a sensitivity of 79.7% at 98% specificity in this sample set and contributed the most to the panel. The 98% specificity criterion was chosen so that no more than 2% of women will be referred to TVS in the next phase of a two-step screening strategy. CA 72-4 could be replaced by CA19-9 or CA15-3 without a reduction of classification power, if necessary. However, a single time point preoperative serum measurement, as was utilized in this study, is arguably not sufficient to evaluate the potential of the biomarker panel and evaluation in longitudinal pre-clinical sera of women destined to develop ovarian cancer is needed. Also, this will help assess if any of these biomarkers provide additional lead times prior to CA125 elevation and also, if they can identify cases that are missed by CA125 alone. While a comprehensive multi-institutional study conducted by the EDRN (Early Detection Research Network) using the PLCO (Prostate Lung Colorectal and Ovarian Cancer Screening Trial) study samples recently questioned the utility of biomarker panels in pre-clinical sera and found that the other promising markers do not add to the power of CA125 to detect early stage ovarian cancer, these studies have utilized samples from a single time-point and do not take advantage of longitudinal biomarker behavior which facilitated the promising results in the UKCTOCS and NROSS trial with CA125 [15]. Moreover, CA125 and HE4 have shown lead times at least close to a year in pre-clinical samples assessed longitudinally from the Carotene and Retinol Trial [16]. The PLCO study utilized single cut-off values of CA125 rather than longitudinal values in the UKCTOCS and NROSS for screening, which could have resulted in not realizing the benefits from early detection. The early stage specimens utilized in this study, obtained from the GOG, do not comprise longitudinal specimens and hence such analysis was infeasible in this study.

Akin to the single marker ROCA, an algorithm that incorporates longitudinal values of a biomarker will require two statistical models: one that models stable biomarker levels in healthy women and one that models one or more markers rising significantly above baseline levels in women that are subsequently diagnosed with ovarian cancer. In our study, we assessed the longitudinal levels for each biomarker utilizing serial samples from apparently healthy volunteers. All biomarkers had a relatively low within-person CV compared with the between-person CV (with the exception of HE4 where they were similar) indicating that each woman has her own baseline for each biomarker, an important criterion for these biomarkers to have utility when interpreted longitudinally. Since there are only a few longitudinal data points per woman, estimation of variation over time within a woman will be imprecise. A hierarchical model describing the distribution of biomarker values over the population of women without ovarian cancer substantially improves the efficiency of estimation as it borrows information across volunteers. Such a hierarchical model

incorporates the distributions of the baseline between-women and within-woman variation over time. The log-transformation for all candidates was the best transformation to achieve a Gaussian distribution except for CA72.4 which required a $\log(\log(.)+1)$ transformation. This study established the appropriate scales for each marker enabling rapid development of modeling of these markers when measured in future studies. Choice of optimal transformations is important as Gaussian models if incorrectly applied to non-symmetric distributions, can result in substantial bias, offsetting the gains that will be derived from the efficiency of borrowing across subjects.

The optimal panel from this study is now under validation in a large blinded retrospective study utilizing longitudinal cases and controls specimens from the UKCTOCS trial, to establish final parameters for the MROCA and to assess both lead times and CA125 complementarity. If either lead time or complementarity advantage is noted for the MROCA over the CA125 ROCA, a prospective trial will be conducted to validate the multi-marker panel, interpreted with the MROCA algorithm to determine clinical utility for this screening methodology.

Acknowledgments

These studies were supported in part by a grant from the MD Anderson SPORE in Ovarian Cancer NCI P50 CA83639 grant; the Shared Resources of the MD Anderson CCSG NCI P30 CA16672 grant; the Cancer Prevention and Research Institute of Texas grant RP101382; the Ovarian Cancer Research Fund; the National Foundation for Cancer Research; and philanthropic support from the Zarrow Foundation and Stuart and Gaye Lynn Zarrow, Golfers Against Cancer, the Kaye Yow Foundation, and the Mossy Family Foundation. CA 72-4 reagents were provided by Roche Diagnostics.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin.* 2016
2. Siegel R, et al. Cancer statistics, 2014. *CA Cancer J Clin.* 64(1):9–29. [PubMed: 24399786]
3. Badgwell D, Bast RC Jr. Early detection of ovarian cancer. *Dis Markers.* 2007; 23(5–6):397–410. [PubMed: 18057523]
4. Jacobs I, et al. Prevalence screening for ovarian cancer in postmenopausal women by CA 125 measurement and ultrasonography. *BMJ.* 1993; 306(6884):1030–1034. [PubMed: 8490497]
5. Buys SS, et al. Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. *JAMA.* 2011; 305(22):2295–2303. [PubMed: 21642681]
6. Skates SJ, et al. Calculation of the risk of ovarian cancer from serial CA-125 values for preclinical detection in postmenopausal women. *J Clin Oncol.* 2003; 21(10 Suppl):206s–210s. [PubMed: 12743136]
7. Menon U, 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(4):327–340. [PubMed: 19282241]
8. Lu KH, et al. A 2-stage ovarian cancer screening strategy using the Risk of Ovarian Cancer Algorithm (ROCA) identifies early-stage incident cancers and demonstrates high positive predictive value. *Cancer.* 2013; 119(19):3454–3461. [PubMed: 23983047]
9. Jacobs IJ, et al. Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet.* 2015
10. Rosen DG, et al. Potential markers that complement expression of CA125 in epithelial ovarian cancer. *Gynecol Oncol.* 2005; 99(2):267–277. [PubMed: 16061277]

11. Yurkovetsky ZR, et al. Multiple biomarker panels for early detection of ovarian cancer. *Future Oncol.* 2006; 2(6):733–741. [PubMed: 17155900]
12. Yurkovetsky Z, et al. Development of a multimarker assay for early detection of ovarian cancer. *J Clin Oncol.* 2010; 28(13):2159–2166. [PubMed: 20368574]
13. Lunn, David J.; Nicky Best, AT.; Spiegelhalter, David. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing.* 2000; 10(4):325–337.
14. Moore RG, et al. The use of multiple novel tumor biomarkers for the detection of ovarian carcinoma in patients with a pelvic mass. *Gynecol Oncol.* 2008; 108(2):402–408. [PubMed: 18061248]
15. Cramer DW, et al. Ovarian cancer biomarker performance in prostate, lung, colorectal, and ovarian cancer screening trial specimens. *Cancer Prev Res (Phila).* 2011; 4(3):365–374. [PubMed: 21372036]
16. Anderson GL, et al. Assessing lead time of selected ovarian cancer biomarkers: a nested case-control study. *J Natl Cancer Inst.* 2010; 102(1):26–38. [PubMed: 20042715]

Figure 1. Dot plots of Log₂-transformed individual biomarker concentrations for the evaluated biomarkers along with the suggested cut-off for the highest sensitivity at a given specificity for the cohort under assessment.

Figure 2. Distribution of within-person variance across healthy volunteers for CA125, HE4 and MMP-7 follows a normal distribution whereas a log-transform is required for CA 72-4 to follow a normal distribution. The blue line represents the median value.

Table 1

Characteristics of the patient population

Histology	No. of Patients	Age (Years)		No. of Longitudinal samples		Race (No. of Patients)					
		Median	Mean	Median	Mean	W	B	H	I	A	O
Healthy	217	65	66.8	5	4.9	193	6	12		6	
Postmenopausal											
Ovarian Cancer	142	54	55			125	6	1	2		8
Stages IA-IC											
Endometrioid	55										
Mucinous	33										
Serous	28										
Clear Cell	19										
Mixed	5										
Undifferentiated	2										

W-White B-Black H-Hispanic I-American Indian A-Asian O-Other

Table 2

Performance of multi-marker panels assessed with a linear classifier algorithm in the training and test sets at 98% specificity.

Marker Panel				Sensitivity at 98% specificity			
M1	M2	M3	M4	Training Mean	Test Mean	Training S.D.	Test S.D.
CA125	CA72-4	HE4	sVCAM	83.3	83.7	4.9	8.3
CA125	CA72-4	HE4	MMP-7	82.1	83.2	3.4	4.4
CA125	CA19-9	HE4	sVCAM	82.1	82.3	3.9	7.8
CA125	CA72-4	HE4		82.9	82.2	3.6	5.1
CA125	MMP-7	HE4	sVCAM	85.5	82.2	3.4	6.2
CA125	CA72-4	HE4	CA19-9	83.4	81.8	3.0	4.6
CA125	CA72-4	MMP-7	sVCAM	85.1	81.6	5.0	7.0
CA125	CA15-3	HE4	sVCAM	82.8	81.5	3.9	6.5
CA125	CA72-4	HE4	CEA	84.0	81.3	4.4	6.6
CA125	CA72-4	HE4	CA15-3	83.3	80.7	4.0	6.4
CA125	CEA	HE4	sVCAM	83.1	80.5	3.1	6.0
CA125	CEA	MMP-7	sVCAM	83.2	79.6	4.4	7.0

Table 3

Within-person variation (%CV) of individual markers assessed in longitudinal healthy controls.

Marker	Mean	25%Qtl	Median	75%Qtl
CA15-3	13.7	8.3	12.3	18.4
CA125	16.7	10.2	14.9	20.1
CA19-9	18.0	6.8	12.3	19.8
CA72-4	30.6	8.6	21.0	46.2
HE4	26.5	19.3	24.9	31.4
MMP-7	21.5	13.0	20.2	27.7
sVCAM	23.1	14.1	22.8	30.6
CEA	24.8	15.6	23.2	31.4

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4
Between-person variation of individual markers assessed in longitudinal healthy controls.

Marker	CA125	CA724	log(CA724)	HE4	MMP7
CV baseline	49%	84%	48%	20%	35%
Median within CV	15%	21%	15%	25%	21%
Within/Between	0.31	0.25	0.31	1.25	0.60