



Use of a Symptom Index, CA125, and HE4 to predict ovarian cancer

M. Robyn Andersen ^{a,b,*}, Barbara A. Goff ^{c,d}, Kimberly A. Lowe ^{a,e}, Nathalie Scholler ^{a,f}, Lindsay Bergan ^a, Charles W. Drescher ^{a,g}, Pamela Paley ^{a,g}, Nicole Urban ^{a,b}

^a Molecular Diagnostics Program, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA

^b School of Public Health and Community Medicine, University of Washington, Seattle, Washington, USA

^c School of Medicine, University of Washington, Seattle, Washington, USA

^d Seattle Cancer Care Alliance, Seattle, Washington, USA

^e Exponent Health Sciences, Seattle, Washington, USA

^f University of Pennsylvania, Philadelphia, Pennsylvania, USA

^g Pacific Gynecology Specialists, Seattle, Washington, USA

ARTICLE INFO

Article history:

Received 10 September 2009

Available online 28 November 2009

Keywords:

Ovarian cancer

CA125

HE4

Symptom Index

Sensitivity

Specificity

ABSTRACT

Background. Prior studies suggest that combining the Symptom Index (SI) with a serum HE4 test or a CA125 test may improve prediction of ovarian cancer. However, these three tests have not been evaluated in combination.

Methods. A prospective case–control study design including 74 women with ovarian cancer and 137 healthy women was used with logistic regression analysis to evaluate the independent contributions of HE4 and CA125, and the SI to predict ovarian cancer status in a multivariate model. The diagnostic performance of various decision rules for combinations of these tests was assessed to evaluate potential use in predicting ovarian cancer.

Results. The SI, HE4, and CA125 all made significant independent contributions to ovarian cancer prediction. A decision rule based on any one of the three tests being positive had a sensitivity of 95% with specificity of 80%. A rule based on any two of the three tests being positive had a sensitivity of 84% with a specificity of 98.5%. The SI alone had sensitivity of 64% with specificity of 88%. If the SI index is used to select women for CA125 and HE4 testing, specificity is 98.5% and sensitivity is 58% using the 2-of-3-positive decision rule.

Conclusions. A 2-of-3-positive decision rule yields acceptable specificity, and higher sensitivity when all 3 tests are performed than when the SI is used to select women for screening by CA125 and HE4. If positive predictive value is a high priority, testing by CA125 and HE4 prior to imaging may be warranted for women with ovarian cancer symptoms.

© 2009 Elsevier Inc. All rights reserved.

Introduction

Ovarian cancer is the second most commonly diagnosed gynecologic malignancy in the United States; it is also the most deadly because over 70% of women with ovarian cancer are diagnosed with advanced stage disease when cure rates are only 20–30% [1]. Ovarian cancer meets the World Health Organization's criteria for a disease that would benefit from screening [2]. However, because current screening modalities have not been shown to reduce the morbidity or mortality of this disease [3], the National Institutes of Health (NIH) Consensus Panel on Ovarian Cancer currently recommends screening only for women at elevated-risk of disease due to a family history [4]. Thus at this time most diagnoses of ovarian cancer start with

evaluation of women's spontaneous complaints of suspicious symptoms or as a result of tests such as ultrasounds conducted for other reasons.

Finding a screening test for ovarian cancer is challenging because ovarian cancer is not a common disease [5]. High-risk women can be identified who are more likely to benefit from intensive screening than average-risk women, but only 10% of ovarian cancer occurs in these women [5]. Multi-modal screening of women at high-risk for ovarian cancer using CA125 and transvaginal sonography (TVS) is recommended for those at highest risk, and is being studied in large efficacy trials in average-risk post-menopausal women [6,7]. When used as a first-line screen, TVS may be sensitive but produces a relatively high rate of false positive results and a potentially unacceptable number of surgeries per cancer found [8]. The use of CA125 as a first-line screen to select women for imaging by TVS as a second-line screen is a promising approach [9], but it has been reported that CA125 is elevated above reference levels in only 50% of clinically detectable early stage patients [3,10,11]. Efforts are underway to improve the performance of CA125 [12], and to identify

* Corresponding author. Translational Outcomes Research Group, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N. M2-B230, PO Box 19024, Seattle, Washington 98109-1024, USA. Fax: +1 206 667 7264.

E-mail address: rande@fhcrc.org (M.R. Andersen).

additional biomarkers for ovarian cancer [13–15]. The use of novel markers in a screening strategy is also being explored (NIH/NCI Grant P50 CA083636). These strategies use imaging prior to surgery to confirm the existence of a mass, and thus may be limited by the sensitivity of imaging.

One of the most promising new serum biomarkers is human epididymis protein 4 (HE4) [5]. HE4 (gene name *WFDC2*) is a glycoprotein that is highly expressed by ovarian carcinomas [16] and [17] its highest normal tissue expression is in the trachea and salivary gland [18]. It has been proposed as a potential biomarker for ovarian cancer as it is expressed by 32% of ovarian cancers without CA125 expression, and, in combination with CA125, serum HE4 has been shown to improve prediction of malignancy in ovarian masses [19]. HE4 was recently approved by the Food and Drug Administration (FDA) for use in the U.S. to monitor ovarian cancer patients for disease recurrence. We have shown that the diagnostic accuracy of HE4 to differentiate cases from healthy controls is similar in high-risk and average-risk women (AUC = 0.931 and AUC = 0.928, respectively, $p = 0.94$) [20].

Until recently symptoms of ovarian cancer were thought to develop only after the disease had progressed to an advanced stage, but now it is appreciated that women with early stage disease often report non-specific symptoms. Symptoms that are new to an individual and occur frequently may distinguish cancer cases from healthy women [21–23] and may be useful in identifying women with ovarian cancer for diagnostic testing; questions assessing these conditions could be used proactively by physicians routinely as a screening tool. We have developed a Symptom Index (SI) that yields a decision rule that might be employed for either purpose. Women reporting pelvic or abdominal pain, bloating, increased abdominal size, difficulty eating or feeling full quickly more than 12 times per month and reporting that these symptoms have occurred for less than one year are considered to have a positive SI [23]. When symptoms are proactively solicited, only 2% of women in a general clinic sample reported symptoms consistent with a positive SI score [23], and a study of its prospective use in a clinic is currently underway. In a sample of high-risk women with pedigrees consistent with inherited susceptibility, 10% reported symptoms consistent with a positive SI [24].

Our prior reports have described the development of the SI [23] and explored its potential use in combination with a CA125 blood test [24]. We reported that symptoms consistent with a positive SI occurred in half of women with ovarian cancer who had normal CA125 levels at a blood draw immediately prior to their diagnostic surgery, and that 53% of women with ovarian cancer were positive on both the SI and CA125 [24]. Here we report the sensitivity and specificity of decision rules using HE4, CA125, and the SI as potential components of a multi-modal multi-step screening program.

Methods

Study population

The study population includes 74 women with ovarian cancer and 137 healthy screening controls. Cases and controls completed identical surveys asking about the frequency (number of days per month) and the duration (number of months) of symptoms that may be associated with ovarian cancer [21]. Cases were surveyed prior to surgery and before receiving a definitive diagnosis of ovarian cancer. Cases were enrolled in the Surgical Specimen Donation protocol of the Pacific Ovarian Cancer Research Consortium (POCRC) and donated a blood sample prior to surgery either at the pre-operative clinic visit or on the day of surgery in the operating room, while under anesthesia and prior to the incision. Of the 74 cases, 14 (19%) have a pedigree consistent with inherited

susceptibility. Controls completed surveys as part of ovarian cancer screening visits conducted on a quarterly basis prior to the collection of blood or performance of TVS. The symptom survey that was completed immediately prior to the blood collection from which biomarker levels were measured was selected for analysis. Controls were enrolled in the POCRC Ovarian Cancer Early Detection Study (OCEDS); all have family histories consistent with inherited susceptibility suggesting high risk for ovarian cancer. The eligibility criteria for OCEDS have been previously described [25]. As only one case reported a documented mutation in BRCA, she was excluded from the analysis as were controls who reported a documented mutation. All women provided informed consent. Approval for this study was obtained from the Institutional Review Board of the Fred Hutchinson Cancer Research Center and area hospitals with participating patients.

Determination of ovarian cancer stage at diagnosis

Whenever possible, women with ovarian malignancies were comprehensively staged by a gynecological oncologist. Women with ovarian cancer diagnosed at stages 1 or 2 were considered to have early stage disease ($n = 31$). Those diagnosed at stages 3 or 4 were considered to have late stage disease ($n = 41$). Two women who were unstaged were not included in stage-specific analyses.

Blood collection and processing

Blood was collected from both cases and controls according to the standard POCRC research protocol. Blood samples sat at room temperature for at least 30 min after collection and before processing to allow clotting. Samples were centrifuged at $1200 \times g$ for 10 min. The serum was then collected and stored at -80°C until analyzed.

Measurement of CA125 and HE4

CA125 and HE4 were measured in serum by sandwich ELISA on a Luminex platform without multiplexing using monoclonal antibodies. CA125 and HE4 serum levels were assessed using bead-based immunoassays performed as described by Scholler et al. [14]. Briefly, bead-based assays were carried out in 96 well MultiScreen®GV filter plates (Millipore Corporation, Billerica, MA) using a vacuum manifold (Millipore) to drain assay reagents. Plates were analyzed with the Bio-Plex Array reader (Bio-Rad, Nercules, CA).

For the CA125 assay, complementary anti-CA125 mouse monoclonal antibodies (mAb) X306 and X52 were purchased from Research Diagnostics, Inc. (RDI) (Flanders, NJ). The CA125 bead-based assay yields values that are strongly correlated ($r = 0.95$) with the research standard CA125II RIA from Fujirebio Diagnostics, Inc. (FDI, Malvern, PA) [14]. For the HE4 assay, complementary anti-HE4 mAbs 3D8 and 2H5 were kind gifts from Dr. Ingegerd Hellstrom [16]. The assay was performed as described by Scholler et al. [14] with the following modifications: antibody-coated beads were incubated with 10-fold diluted sera and captured antigens were detected with $2 \mu\text{g/ml}$ of biotinylated 3D8. The HE4 bead-based assay yields values that are strongly correlated ($r = 0.90$) with a plate-based assay using the same antibodies [14].

Thresholds for a positive CA125, HE4, and Symptoms Index

The thresholds for positivity of CA125 and HE4 were determined by dichotomizing each marker at the 95th percentile in the control group. Cases and controls with marker levels above this threshold were considered to have a positive test result; all others were considered to have a negative test result. Consistent with prior reports [23], women were classified as having a positive SI if they reported bloating or increased abdominal size, abdominal or pelvic pain, or

Screening Tests	Description
Single Markers	
CA125	CA125 was dichotomized at the 95 th percentile in the control group. Subjects with a marker value above that threshold were considered to be positive for CA125.
HE4	HE4 was dichotomized at the 95 th percentile in the control group. Subjects with a marker value above that threshold were considered to be positive for HE4.
Symptom Index	The Symptom Index was considered to be positive if the patient had at least one of the following symptoms for less than one year but more than 12 times per month: bloating or increased abdominal size, abdominal or pelvic pain, difficulty eating or feeling full quickly.
Marker Combinations	
CA125 or HE4	Screen considered positive if CA125, HE4 or both were positive.
Symptom Index or CA125	Screen considered positive if either the Symptoms Index or CA125 was positive, or if both were positive.
Symptom Index or HE4	Screen considered positive if either the Symptoms Index or HE4 was positive, or if both were positive.
Any 1 of 3 tests positive	Screen considered positive if any one of the Symptoms Index or CA125 or HE4 was positive, or if two or more tests were positive.
Any 2 of 3 tests positive	Screen considered positive if both the Symptoms Index and CA125 were positive, or if both the Symptoms Index and HE4 were positive, or if both CA125 and HE4 were positive, or if all three tests were positive.
Symptom Index and at least 1 additional test positive	Classified as positive if the Symptom Index was positive <i>in addition to</i> either a positive CA125 or a positive HE4, or if all three tests were positive.

Fig. 1. Description of screening tests and biomarker combinations.

difficulty eating or feeling full quickly more than 12 times per month and occurring newly within the past 12 months.

Statistical analysis

The cases and controls were frequency matched on age above or below 50 prior to analysis. STATA statistical software package (version 10.0, Stata Corporation, College Station, TX) was used for unconditional logistic regression analysis to determine if each test (the SI, CA125, and HE4) independently predicted cancer after controlling for the contribution of the other two tests. All statistical tests were two-sided and considered to be statistically significant at $p < 0.05$. Baseline age was dichotomized at 50 years. The age, stage, SI, and dichotomized CA125 and HE4 levels were compared across the study populations using the Fisher's exact test.

The diagnostic accuracy of one-, two-, and three-marker decision rules were determined by calculating the sensitivity (defined as true positives divided by the sum of true positives and false negatives) and specificity (defined as true negatives divided by the sum of true negatives and false positives) of the overall test. We explored alternative decision rules for using the three markers including requiring that 2 of the 3 markers be positive to call the overall test positive, and requiring that the 2 positive tests include the SI. The sensitivity and specificity of the tests and their combinations were also calculated for women age 50 and over and those under the age of 50, for those with early and late stage disease, and for high-risk women only.

Results

The 74 cases from the surgical population included 6 women with mucinous cancer, 6 women with clear cell carcinoma, 7 women with endometrioid cancer, 5 women with other adenocarcinomas, and 50 women with serous cancer. Cases were more likely to be positive for the SI, CA125, and HE4 ($p < 0.001$ for each marker) than were the 137 controls from the screening population. CA125, HE4, and the SI were all highly significant predictors of ovarian cancer in the logistic regression model ($p < 0.001$ for each marker), which explained 68% of the variability in case status.

Fig. 1 defines the marker combination decision rules that were evaluated. Table 1 reports the numbers of cases and controls by age

Table 1

Numbers of cases and controls by age group, risk status, stage (cases only), and screen positivity by CA125, HE4, Symptom Index, and marker combinations.

	Controls, n = 137	Cases, n = 74	p- value
Age ^a			
<50 years	22 (16%)	12 (16%)	0.98
>50 years	109 (80%)	59 (80%)	
Risk status			
Average	0	60 (81%)	NA
High	137 (100%)	14 (19%)	
Stage ^b			
Early	NA	31 (42%)	NA
Late	NA	41 (55%)	
CA125 ^c			
Positive	7 (5%)	60 (81%)	<0.001
Negative	130 (95%)	14 (19%)	
HE4 ^c			
Positive	7 (5%)	57 (77%)	<0.001
Negative	130 (95%)	17 (23%)	
Symptom Index ^c			
Positive	16 (12%)	47 (64%)	<0.001
Negative	121 (88%)	27 (36%)	
CA125 or HE4 ^c			
Positive	14 (10%)	66 (89%)	<0.001
Negative	123 (90%)	8 (11%)	
Symptom Index or CA125 ^c			
Positive	23 (17%)	68 (92%)	<0.001
Negative	114 (83%)	6 (8%)	
Symptom Index or HE4 ^c			
Positive	21 (15%)	68 (92%)	<0.001
Negative	116 (85%)	6 (8%)	
Any 1 of 3 tests positive ^c			
Positive	28 (20%)	70 (95%)	<0.001
Negative	109 (80%)	4 (5%)	
Any 2 of 3 tests positive ^c			
Positive	2 (2%)	62 (84%)	<0.001
Negative	135 (98%)	12 (16%)	
Symptom Index and at least 1 additional test positive ^c			
Positive	2 (2%)	43 (58%)	<0.001
Negative	135 (98%)	31 (42%)	

^a Data missing for 6 subjects in the control group and 3 cases.

^b Data missing for 2 cases.

^c See Fig. 1 for descriptions of the screening tests and biomarker combinations.

Table 2

Diagnostic accuracy and 95% CI for CA125, HE4, the Symptom Index, and marker combinations by age group.

Screening tests ^a	Number of cases with positive screening test	Sensitivity ^b	Number of controls with negative screening test	Specificity ^c
	Total: n = 74; <50 years old: n = 12; >50 years old: n = 59		Total: n = 137; <50 years old: n = 22; >50 years old: n = 109	
CA125	60	81.1 (70.3–89.3)	130	94.9 (89.9–97.9)
<50 years	10	83.3 (51.6–97.9)	20	90.9 (70.8–98.9)
>50 years	48	81.4 (69.1–90.3)	105	96.3 (90.9–99.0)
HE4	57	77.0 (65.8–86.0)	130	94.9 (89.8–97.9)
<50 years	10	83.3 (51.6–97.9)	22	100 (84.6–100)
>50 years	45	76.3 (63.4–86.4)	102	93.6 (87.2–97.4)
Symptom Index (SI)	47	63.5 (51.5–74.4)	121	88.3 (81.7–93.2)
<50 years	10	83.3 (51.5–97.9)	20	90.0 (70.8–98.9)
>50 years	34	57.6 (44.1–70.4)	97	89.0 (81.6–94.2)
CA125 or HE4	66	89.2 (79.8–95.2)	123	89.8 (83.4–94.3)
<50 years	12	100 (73.5–100)	20	90.9 (70.8–98.9)
>50 years	52	88.1 (77.1–95.1)	98	89.9 (82.7–94.9)
SI or CA125	68	91.9 (83.2–97.0)	114	83.2 (75.9–89.0)
<50 years	11	91.7 (61.5–99.8)	18	81.8 (59.7–94.8)
>50 years	54	91.5 (81.3–97.2)	93	85.3 (77.3–91.4)
SI or HE4	68	91.9 (83.2–97.0)	116	84.7 (77.5–90.3)
<50 years	12	100 (73.5–100)	20	90.9 (70.8–98.9)
>50 years	53	89.8 (79.2–96.2)	92	84.4 (76.2–90.6)
Any 1 of 3 tests positive	70	94.6 (86.7–98.5)	109	79.6 (71.8–86.0)
<50 years	12	100 (73.5–100)	18	81.8 (59.7–94.8)
>50 years	55	93.2 (83.5–98.1)	88	80.7 (72.1–87.7)
Any 2 of 3 tests positive	62	83.8 (73.4–91.3)	135	98.5 (94.8–99.8)
<50 years	11	91.7 (61.5–99.8)	22	100 (84.6–100)
>50 years	49	83.1 (71.0–91.6)	107	98.2 (93.5–99.8)
SI and at least 1 additional test positive	43	58.1 (46.1–69.5)	135	98.5 (94.8–99.8)
<50 years	10	83.3 (51.6–97.9)	22	100 (84.6–100)
>50 years	31	52.5 (39.1–65.7)	107	98.2 (93.5–99.8)

^a See Table 1 for descriptions of the screening tests and biomarker combinations.

^b Sensitivity = true positives / (true positives + false negatives).

^c Specificity = true negatives / (true negatives + false positives).

group, risk status, stage (cases only), and screen positivity by CA125, HE4, SI, and marker combinations, summarizing the overall sensitivity and specificity of each decision rule. Table 2 provides estimates of sensitivity and specificity within age (above or below age 50) and stage (early stage vs. late stage). Table 3 provides estimates of sensitivity and specificity for average-risk and high-risk cases and for early-stage and late-stage cases separately.

As a single marker, CA125 had the highest overall sensitivity at 95% specificity, identifying 81.1% of the case population overall including 78.6% of the high-risk cases and 67.7% of the early stage cases. HE4 had the highest overall sensitivity at 95% specificity in high-risk cases, identifying 100% of the 14 high-risk cases. The SI alone yielded sensitivity of 63.5% and specificity of 88.3%. The two-marker decision rules combining the SI with either CA125 or HE4 identified 91.9% of the cases overall and 100% of both the early-stage cases and the high-risk cases, with specificity of 83% to 85% for CA125 and HE4, respectively. Better specificity was achieved using CA125 and HE4 without the SI, identifying 89.2% of the cases overall at a specificity of 89.8%. A three-marker decision rule defining a test positive if any one of CA125, HE4, or the SI was positive identified 94.6%, 90.3%, and 100% of the overall, early-stage and high-risk cases, respectively, with a false positive rate of 20.4% (specificity 79.6%).

The three-marker decision rule yields high sensitivity but poor specificity; in general, there is a trade-off between sensitivity and specificity. Specificity can be improved by requiring that 2 of the 3 tests be positive. A three-marker decision rule requiring 2 of 3 tests to be positive identifies 83.8%, 67.7%, and 100% of the overall, early-stage, and high-risk cases, respectively, with a false positive rate of 1.5% (specificity 98.5%). Requiring that one of the 2-of-3 positive markers is the SI identified 58.1%, 35.5%, and 64.3% of the overall, early-stage, and high-risk cases, respectively, with the same false positive rate of 1.5% (specificity 98.5%).

Discussion

Prior studies suggest that combining either a proactively solicited SI [24] or a serum HE4 test [26] with a serum CA125 test could improve ovarian cancer diagnostic test performance. We evaluated all three tests in combination using a prospective case-control study design including 74 women with ovarian cancer and 137 healthy women. Logistic regression analysis was used to evaluate the independent contributions of HE4, CA125, and the SI to the prediction of ovarian cancer status in a multivariate model; each test made a significant independent contribution. We also assessed the diagnostic performance of various decision rules using combinations of these tests.

In this study, HE4 and CA125 performed similarly overall both on their own and in combination with the SI. HE4 performed somewhat better (100% vs. 78.6% sensitivity at 95% specificity for HE4 and CA125, respectively) in high-risk women, the population of greatest interest because it is the only group for whom screening is currently recommended. This finding is consistent with recent reports that HE4 outperforms CA125 as a first-line screen due to its high sensitivity [27]. Specificity can be controlled by choice of a threshold for CA125 and HE4, but for the SI, it depends on women's symptom reports. In this study, specificity of the SI was 88.3%, consistent with previous reports of symptom reporting among high-risk women [23,24]. Improvement in specificity was achieved by a decision rule requiring at least 2 of the 3 tests to be positive. When requiring that one of the 2 positive tests be the SI (blood work performed only in women with symptoms), this rule had specificity of 98.5% with overall sensitivity of 58.1% (35.5% and 64.3% for early-stage disease and high-risk women, respectively). Use of all three tests together as a first-line screen improves sensitivity dramatically while preserving specificity: This rule had excellent overall sensitivity of 83.8% (67.7% and 100% for early-stage disease and

Table 3
Sensitivity of the screening tests by risk status and ovarian cancer stage.

Screening tests ^b	Sensitivity by risk status ^a				Sensitivity by ovarian cancer stage ^a			
	Number of average-risk cases with positive screening test (n = 60)	Sensitivity in average-risk cases	Number of high-risk cases with positive screening test (n = 14)	Sensitivity in high-risk cases	Number of early stage cases with positive screening test (n = 31)	Sensitivity in early stage cases	Number of late stage cases with positive screening test (n = 41)	Sensitivity in late stage cases
CA125	49	81.7 (69.6–90.5)	11	78.6 (49.2–95.3)	21	67.7 (48.6–83.3)	38	92.7 (80.1–98.5)
HE4	43	71.7 (58.6–82.5)	14	100 (76.8–100)	18	58.1 (39.1–75.5)	38	92.7 (80.1–98.5)
Symptom Index (SI)	38	63.3 (49.9–75.4)	9	64.3 (35.1–87.2)	14	45.2 (27.3–64.0)	32	78.0 (62.4–89.4)
CA125 or HE4	52	86.7 (75.4–94.1)	14	100 (76.8–100)	25	80.6 (62.5–92.5)	40	97.6 (87.1–99.9)
SI or CA125	54	90.0 (79.5–96.2)	14	100 (76.8–100)	26	83.9 (66.3–94.5)	40	97.6 (87.1–99.9)
SI or HE4	54	90.0 (79.5–96.2)	14	100 (76.8–100)	26	83.9 (66.3–94.5)	40	97.6 (87.1–99.9)
Any 1 of 3 tests positive	56	93.3 (83.8–98.2)	14	100 (76.8–100)	28	90.3 (74.2–98.0)	40	97.6 (87.1–99.9)
Any 2 of 3 tests positive	48	80.0 (67.7–89.2)	14	100 (76.8–100)	21	67.7 (48.6–83.3)	40	97.6 (87.1–99.9)
SI and at least 1 additional test positive	34	56.7 (43.2–69.4)	9	64.3 (35.1–87.2)	11	35.5 (19.2–54.6)	32	78.0 (62.4–89.4)

^a Sensitivity = true positives/(true positives + false negatives).

^b See Table 1 for descriptions of the screening tests and biomarker combinations.

high-risk women, respectively) and specificity of 98.5%, suggesting its potential utility as a first-line screen to select women for imaging.

In the absence of periodic proactive screening, average-risk women may report symptoms to their physicians. No longer considered to be a “silent killer,” ovarian cancer now is believed to cause non-specific abdominal symptoms, and recent guidelines alert women and physicians specifically to new, frequent symptoms. The American Cancer Society now recommends women to see their doctor if they experience symptoms of abdominal swelling or bloating, pelvic pressure or pain, difficult eating or feeling full quickly, or problems with urination [28]. We used women’s solicited self-report of symptoms associated with bloating, pain, and difficulty eating that were experienced frequently (more than 12 days a month) and newly (beginning within the past 12 months) to create our SI [23]; alternative approaches to defining relevant symptoms have also been described [29].

Appropriate follow-up when symptoms are reported, either as spontaneous complaints or because they have been proactively solicited by a physician, has not yet been defined. In either case, physicians might refer women with symptoms for pelvic imaging to confirm that the symptoms are associated with an ovarian mass. However, referral for imaging of all women with symptoms is likely to identify many more benign than malignant conditions [29] and lead to unnecessary surgery. An alternative approach might be to test such women for elevation of CA125 and/or HE4 prior to referral for ultrasound. This approach could potentially reduce the need for imaging as well as unnecessary surgery in the follow-up of women’s reports of nonspecific symptoms. In this analysis, 91.5% of women with cancer who reported symptoms were positive on either CA125 or HE4. Thus, this approach identified all but 4 (8.5%) of the 48 ovarian cancer cases identified by symptoms reports alone, yielding overall detection of 58% of the women with ovarian cancer. High specificity was obtained, with a low rate of false positive results (1.5% of the control group).

The strengths of our study include the prospective collection of data on 3 test modalities that have not been evaluated in combination previously. A limitation of our study is that women included in the control group all have a family history of ovarian and/or breast cancer whereas only 19% of the cases have such a pedigree. The design, driven by the fact that we screen only women with a family history,

may bias our results conservatively because symptoms are more frequently reported by the high-risk population than by the general clinic population; specificity of the SI would have been better if we had used controls matched for risk status. Less bias is expected for the markers because recent studies show that the performance characteristics of CA125 and HE4 are largely unaffected by risk status of the population [20]. Offsetting this conservative bias in our study is the potential for recall bias among the cases: women scheduled for surgery due to suspected ovarian cancer may have been even more likely to remember recent symptoms than were the high-risk controls. Because it is difficult to ascertain the degree of these biases, efforts are currently underway to prospectively collect symptoms data in a primary care clinic. A final limitation is that we do not have detailed information for imaging results. All women in the case group of this study had positive imaging that led to surgery and this may have influenced their symptoms reports. Whether or not the symptoms they reported on our questionnaires were also recorded in medical records was not evaluated as part of this study effort.

In average-risk women, the incidence of ovarian cancer is 40/100,000. With this relatively low incidence, a screening test would need a specificity of 99.6% to achieve a positive predictive value of 10%. The 2-of-3 rule yielded a first-line screen specificity of 98.5%; the addition of imaging as a second-line screen would likely further improve specificity perhaps bringing the PPV to an acceptable level. The use of a SI, CA125, and HE4 as an annual first-line screen to select women for imaging if any 2 of the 3 tests are positive warrants further study. Its cost-effectiveness relative to alternative multi-modal screening strategies such as using CA125 alone to select women for imaging as in the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) or using both CA125 and imaging annually as in the Prostate, Lung, Colon and Ovary (PLCO) screening trial, is of particular interest.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

Acknowledgments

Supported by a grant from the Marsha Rivkin Center for Ovarian Cancer Research (Seattle, WA) and by National Institutes of Health/National Cancer Institute grant P50 CA083636 to N. Urban (“Pacific

Ovarian Cancer Research Consortium: Specialized Program of Research Excellence in Ovarian Cancer"). We thank Marcia Gaul, Vandana Oza and Kristi Schurman for administrative support, Shelly Hager for database support, and Canary Foundation for generous contributions to the Translational and Outcomes Research laboratory at the Fred Hutchinson Cancer Research Center.

References

- [1] Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. Cancer statistics. *CA Cancer J Clin* 2007;57(2007):43–66.
- [2] Menon U, Jacobs IJ. Ovarian cancer screening in the general population: current status. *Int J Gynecol Cancer* 2001;11:3–6.
- [3] Ozols RF, Rubin SC, Thomas GM, Robboy SJ. Epithelial ovarian cancer. In: Hoskins WJ, Perez CA, Young R, Barakat R, editors. *Principles and practice of gynecologic oncology*. 4th ed. Philadelphia, PA: Lippincott Williams and Wilkins; 2005. p. 895–987.
- [4] National Institutes of Health Consensus Development Panel. Ovarian cancer: screening, treatment and followup; (1994), Document Number 3.
- [5] Bast Jr RC, Badgwell D, Lu Z, Marquez R, Rosen D, Liu J, et al. New tumor markers: CA125 and beyond. *Int J Gynecol Cancer* 2005;15:274–81.
- [6] Jacobs IJ, Skates SJ, MacDonald N, Menon U, Rosenthal AN, Davies AP, et al. Screening for ovarian cancer: a pilot randomized controlled trial. *Lancet* 1999;353:1207–10.
- [7] Prorok PC, Andriole GL, Bresalier RS, Buys SS, Chia D, Crawford ED, et al. Design of the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial. *Control Clin Trials* 2000;21:273S–309S.
- [8] Lacey JVJ, Greene MH, Buys SS, Reding D, Riley TL, Berg CD, et al. Ovarian cancer screening in women with a family history of breast or ovarian cancer. *Obstet Gynecol* 2006;108:1176–84.
- [9] 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.
- [10] Helzlsouer KJ, Bush TL, Alberg AJ, Bass KM, Zacur H, Comstock GW. Prospective study of serum CA-125 levels as markers of ovarian cancer. *Jama* 1993;269:1123–6.
- [11] Mann WJ, Patsner B, Cohen H, Loesch M. Preoperative serum CA-125 levels in patients with surgical stage I invasive ovarian adenocarcinoma. *J Natl Cancer Inst* 1988;80:208–9.
- [12] Menon U, Skates SJ, Lewis S, Rosenthal AN, Rufford B, Sibley K, et al. Prospective study using the risk of ovarian cancer algorithm to screen for ovarian cancer. *J Clin Oncol* 2005;23:7919–26.
- [13] McIntosh MW, Drescher C, Karlan B, Scholler N, Urban N, Hellstrom KE, et al. Combining CA 125 and SMR serum markers for diagnosis and early detection of ovarian carcinoma. *Gynecol Oncol* 2004;95:9–15.
- [14] Scholler N, Crawford M, Sato A, Drescher CW, O'Brian KC, Kiviati N, et al. Bead-based ELISA for validation of ovarian cancer early detection markers. *Clin Cancer Res* 2006;12:2117–24.
- [15] Skates SJ, Horick N, Yu Y, Xu FJ, Berchuck A, Havrilesky LJ, et al. Preoperative sensitivity and specificity for early-stage ovarian cancer when combining cancer antigen CA-125II, CA 15-3, CA 72-4, and macrophage colony-stimulating factor using mixtures of multivariate normal distributions. *J Clin Oncol* 2004;22:4059–66.
- [16] Hellstrom I, Raycraft J, Hayden-Ledbetter M, Ledbetter JA, Schummer M, McIntosh M, et al. The HE4 (WFDC2) protein is a biomarker for ovarian carcinoma. *Cancer Res* 2003;63:3695–700.
- [17] Drapkin R, von Horsten HH, Lin Y, Mok SC, Crum CP, Welch WR, et al. Human epididymis protein 4 (HE4) is a secreted glycoprotein that is overexpressed by serous and endometrioid ovarian carcinomas. *Cancer Res* 2005;65:2162–9.
- [18] Galgano MT, Hampton GM, Frierson HF, Jr., Comprehensive analysis of HE4 expression in normal and malignant human tissues. *Mod Pathol* 2006;19:847–53.
- [19] Rosen DG, Wang L, Atkinson JN, Yu Y, Lu KH, Diamandis EP, et al. Potential markers that complement expression of CA125 in epithelial ovarian cancer. *Gynecol Oncol* 2005;99:267–77.
- [20] Shah CA, Lowe KA, Paley P, Wallace E, Anderson GL, McIntosh MW, et al. Influence of ovarian cancer risk status on the diagnostic performance of the serum biomarkers mesothelin, HE4, and CA125. *Cancer Epidemiol Biomarkers Prev* 2009;18:1365–72.
- [21] Goff BA, Mandel L, Muntz HG, Melancon CH. Ovarian carcinoma diagnosis: results of a national ovarian cancer survey. *Cancer* 2000;89:2068–75.
- [22] Goff BA, Mandel LS, Melancon CH, Muntz HG. Frequency of symptoms of ovarian cancer in women presenting to primary care clinics. *Jama* 2004;291:2705–12.
- [23] Goff BA, Mandel LS, Drescher CW, Urban N, Gough S, Schurman KM, et al. Development of an ovarian cancer symptom index. *Cancer* 2007;109:221–7.
- [24] Andersen MR, Goff BA, Lowe KA, Scholler N, Bergan L, Drescher C, et al. Combining a Symptoms Index with CA125 to improve detection of ovarian cancer. *Cancer* 2008;113:484–9.
- [25] Lowe KA, Shah C, Wallace E, Anderson G, Paley P, McIntosh M, et al. Effects of personal characteristics on serum CA125, mesothelin, and HE4 levels in healthy postmenopausal women at high-risk for ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 2008;17:2480–7.
- [26] Scholler N, Lowe K, Bergan L, Kampani A, Ng V, Forrester R, et al. Use of yeast-secreted in vivo biotinylated recombinant antibodies (biobodies) in bead-based ELISA. *Clin Cancer Res* 2008;14:2647–55.
- [27] Havrilesky LJ, Whitehead CM, Rubatt JM, Cheek RL, Groelke J, He Q, et al. Evaluation of biomarker panels for early stage ovarian cancer detection and monitoring for disease recurrence. *Gynecol Oncol* 2008;110:374–82.
- [28] American Cancer Society. *Cancer Facts and Figures 2008*. Atlanta, GA: American Cancer Society (2008) (updated 2008 cited 2009 Aug 24) Available from, <http://www.cancer.org/downloads/STT/2008CAFFfinalsecured.pdf>.
- [29] Pavlik EJ, Saunders BA, Doran S, McHugh KW, Ueland FR, Desimone CP, et al. The search for meaning—symptoms and transvaginal sonography screening for ovarian cancer: predicting malignancy. *Cancer* 2009;115:3689–98.