

BRIEF COMMUNICATION

Potential Role of HE4 in Multimodal Screening for Epithelial Ovarian Cancer

Nicole Urban, Jason D. Thorpe, Lindsay A. Bergan, Robin M. Forrest, Archana V. Kampani, Nathalie Scholler, Kathy C. O'Briant, Garnet L. Anderson, Daniel W. Cramer, Christine D. Berg, Martin W. McIntosh, Patricia Hartge, Charles W. Drescher

Manuscript received February 25, 2011; revised August 11, 2011; accepted August 15, 2011

Correspondence to: Nicole Urban, ScD, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N., M2-B230, Seattle, WA 98109-1024 (e-mail: nurban@fhcrc.org).

In screening for epithelial ovarian cancer, unnecessary surgery can be reduced by limiting use of transvaginal ultrasound (TVU) to women with increasing CA125 serum levels. Replacing or augmenting TVU with measurement of a serum marker specific for malignancy might further improve screening performance. Serum samples from 112 invasive ovarian cancer patients and 706 matched control subjects from the Prostate, Lung, Colorectal, and Ovarian trial were used to evaluate human epididymis protein 4 (HE4), mesothelin, matrix metalloproteinase 7 (MMP7), SLPI, Spondin2, and insulin-like growth factor binding protein 2 (IGFBP2) for their potential use in screening. TVU results were available for a subset of 84 patients and 516 control subjects used to compare the best marker with TVU. HE4 was found to perform better than TVU as a second-line screen, confirming 27 of 39 cancers with increasing CA125 serum levels compared with 17 cancers confirmed by TVU ($P = .03$). Serum HE4 levels were found to increase with age and smoking status, suggesting that a longitudinal algorithm might improve its performance.

J Natl Cancer Inst 2011;103:1630–1634

A recent report from the Prostate, Lung, Colorectal, and Ovarian (PLCO) trial suggests that screening for epithelial ovarian cancer (EOC) using annual transvaginal ultrasound (TVU) and CA125 measurement leads to unnecessary surgery (1) without reducing mortality (2). A multimodal strategy using increasing CA125 serum levels measured annually to select women for TVU yielded an acceptable positive predictive value of 35% at the initial screen in the UK Collaborative Trial of Ovarian Cancer Screening (3); data on patient mortality have not yet been reported. Poor performance by TVU in efficacy trials suggests that use of a second serum marker in participants with increasing CA125 serum levels may offer advantages as a second-line screening method.

Using proximate samples from 112 women with EOC and 706 matched non-oophorectomized control subjects from the PLCO trial, we evaluated six candidate

serum markers for their elevation in patients with and without increasing CA125 serum levels. CA125, human epididymis protein 4 (HE4), mesothelin, matrix metalloproteinase 7 (MMP7), SLPI, Spondin-2, and insulin-like growth factor binding protein 2 (IGFBP2) previously showed at least 30% sensitivity at 95% specificity in clinical studies (4–7) or in validation studies using preclinical samples (8,9).

PLCO trial design and study population (1) are described in the Supplementary Methods (available online). Previous reports from the PLCO validation study did not analyze TVU results or serial measures of CA125 (9,10). Also unique to this analysis are inclusion of 237 control subjects with a family history of breast or ovarian cancer to improve generalizability to the high-risk population. Written informed consent was obtained from each participant. EOC was defined as ovarian, fallopian tube, and primary peritoneal cancer but excluded

granulosa cell tumors. Assays used for study are described in the Supplementary Methods (available online). Marker concentrations were rescaled, and covariable adjustment was performed using analysis of covariance on the standardized scale (8).

In the PLCO trial, women aged 55–74 at enrollment were screened annually for EOC for 6 years by CA125 and considered positive if CA125 levels were 35 U/mL or higher. Serial CA125 results obtained in the trial were available for all years, but only one proximate serum sample was provided for the measurement of novel markers. A clinically significant increase in CA125 in the proximate sample (“increasing CA125” hereafter) was defined by a personal threshold consistent with 96.9% specificity as determined by the parametric empirical Bayes rule, a longitudinal algorithm (11,12) designed to detect increasing marker levels. Because serum was not stored from blood collected at the fourth screen, the proximate sample was obtained at the third screen for cancers diagnosed between the third and fifth screens (Supplementary Table 3, available online), adversely affecting sensitivity estimates in this study. PLCO study participants were also screened annually by TVU in the first 4 years, resulting in a subset including 84 patients and 516 controls subjects (including 175 family history control subjects) with TVU results available.

Patients and control subjects did not differ in most categories; however, the patients did statistically significantly differ from the control subjects in terms of family history of breast or ovarian cancer (by design, $P < .001$ for all participants and $P < .001$ for the subset of participants with TVU results available) and in history of endometriosis ($P = .029$ for all participants and $P < .012$ for the subset of participants with TVU results available); family history control subjects did not statistically significantly differ from the randomly selected PLCO population control subjects (Supplementary Tables 1 and 2, available online).

Using all 112 patients and 706 control subjects with marker results available, we investigated the associations between CA125, HE4, mesothelin, MMP7, IGFBP2,

Spondin-2, and SLPI marker levels with malignancy, accounting for increasing CA125 to address validity as a second-line screen and simultaneously adjusted for participant characteristics (13,14). Statistically significantly increased levels of every marker were observed for cancers identified by increasing CA125 serum levels (Table 1). The highest average signal was recorded for HE4 and was 4.26 SDs above the mean HE4 measurement in control subjects ($P < .001$), suggesting that HE4 is the best marker of those tested for use as a second-line screening modality. For cancers not detected by increasing CA125, an HE4 signal of 0.495 SDs above the mean in control subjects ($P = .006$) was observed; statistically significant changes in the levels of other markers were not observed in these patients (Table 1).

We also investigated associations between marker levels and each participant characteristic after adjusting for the remaining characteristics (Table 1). The characteristics studied included age at first blood draw (continuous), body mass index (continuous), nonwhite race (yes or no), family history (breast or ovarian cancer, yes or no), oral contraceptive use (≥ 1 year, yes or no), nulliparous (yes or no), history of endometriosis (yes or no), current smoker (yes or no), prior hysterectomy (yes or no), current hormone therapy with intact uterus (yes or no), and current hormone therapy with prior hysterectomy (yes or no). All markers, including CA125, were associated with at least two characteristics, but no markers were associated with family history, history of endometriosis, oral contraceptive use, or nulliparity. An increasing CA125 serum level was the only marker independent of all participant characteristics. HE4 serum levels increased with age and smoking suggesting that analysis using a longitudinal algorithm might further improve performance of HE4 as a screening modality, but serial measures were not available to perform this analysis.

HE4 was evaluated for its potential value in multimodal screening using the subset of 84 patients and 516 control subjects with TVU results available (Table 2). Positivity thresholds were chosen for HE4 and for increasing CA125, which yield 96.9% specificity to be consistent with TVU specificity in this dataset. McNemar

test was used to test the hypothesis that the number of cancers detected by measuring HE4 in women with increasing CA125 is greater than the number of cancers detected by performing TVU. Using increasing CA125 levels to select women for follow-up testing, we found that 27 of 39 cancers with increasing CA125 were confirmed by measuring HE4 levels compared with 17 cancers confirmed by TVU ($P = .03$). Positivity of CA125 may be defined by a single threshold rule rather than by an increase; for example, in the PLCO trial, positivity was defined by a CA125 level of at least 35 U/mL (specificity = 98.8%). Using this rule to select women for a confirmatory test, we found that 26 of the 34 CA125-positive cancers were confirmed by measuring HE4 compared with 16 cancers confirmed by TVU ($P = .02$) (data not shown).

Measurement of HE4 had higher sensitivity in confirming all stages of type 2 EOC (includes grade 3 and 4 tumors of serous, undifferentiated, or adenocarcinoma not otherwise specified histology) and lower sensitivity in confirming early-stage type 1 EOC (all remaining tumors including grades 1 and 2 serous tumors and all clear cell, endometrioid, and mucinous histology tumors) (15) compared with TVU. We found that sensitivity is maximized by both measuring HE4 and performing TVU in the second-line screen and calling the screen positive if either confirmatory test is positive. We found that HE4 also performs better than TVU as a first-line screen. As reported in Table 2, increasing CA125 and HE4 serum levels both had higher sensitivity for EOC than did TVU when used alone as a first-line screen at the same 96.9% specificity, identifying 39 and 30 of 84 cancers respectively compared with 24 cancers identified by TVU.

This study used preclinical samples to retrospectively validate clinical decision rules. There are several limitations to the study, including the lack of access to serial samples and the fact that serum samples were not available from the fourth screen. Also, control subjects who underwent oophorectomy during the trial were excluded from the study, upwardly biasing estimates of specificity for screening modalities that falsely identify disease conditions other than EOC for which

CONTEXT AND CAVEATS

Prior knowledge

Recent studies suggest that screening for epithelial ovarian cancer by annual transvaginal ultrasound (TVU) and CA125 measurement does not confer a survival benefit and may lead to unnecessary surgery for some patients.

Study design

A multimodal strategy combining measurement of CA125 and other serum markers specific for ovarian cancer was investigated as a potential screening strategy and compared with TVU. Serum samples from invasive ovarian cancer patients and matched control subjects from the Prostate, Lung, Colorectal, and Ovarian trial were used to evaluate candidate serum markers for their association with malignancy and increasing CA125.

Contribution

Of the six serum markers investigated, human epididymis protein 4 (HE4) was identified as a potential serum marker with the highest relative increase in serum levels and was associated with malignancy. In a sub-analysis of patients with increasing CA125, HE4 was found to confirm more cancers than TVU as a second-line screen.

Implication

HE4 may be a valid serum marker for use in a multimodal screening strategy for ovarian cancer.

Limitations

Measuring HE4 may not make screening itself effective. Although adding HE4 measurement to a screening strategy may improve specificity, it may also reduce sensitivity. Further studies to determine the effect of detecting HE4-associated tumors on patient outcomes and confirmation of better performance as a first- and second-line screen relative to TVU are needed.

From the Editors

oophorectomy is performed. The final limitation of the study is that the PLCO trial used decades old screening test technology, still in use by others because of a lack of technical advances in the field. Potential bias introduced by these limitations likely leads to conservative conclusions. Use of an automated clinical platform to measure analytes in serum, use of increasing HE4 in decision rules, and use

Table 1. Associations of marker levels with patient status by increasing CA125 and participant characteristics reported at the time of enrollment for 112 patients and 706 control subjects in the Prostate, Lung, Colorectal, and Ovarian trial with CA125 and HE4 results available*

| Participant characteristic and CA125 status† | No. (%) | Mean (SD) | CA125 | | CA125 (PEB) | | HE4 | | Mesothelin | | MMP7 | | IGFBP2 | | Spondin-2 | | SLPI | |
|---|------------|-----------|--------|-------|-------------|-------|--------|-------|------------|-------|--------|-------|--------|-------|-----------|-------|--------|-------|
| | | | Coeff | P | Coeff | P | Coeff | P | Coeff | P | Coeff | P | Coeff | P | Coeff | P | Coeff | P |
| All patients† | 112 (13.7) | NA | 2.642 | <.001 | 3.580 | <.001 | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| Patients with increasing CA125‡ | 48 (5.9) | NA | NA | NA | NA | NA | 4.260 | <.001 | 1.211 | <.001 | 0.786 | <.001 | 0.347 | .023 | 0.478 | .001 | 0.845 | <.001 |
| Patients without increasing CA125‡ | 64 (7.8) | NA | NA | NA | NA | NA | 0.495 | .006 | 0.122 | .423 | -0.026 | .765 | 0.025 | .854 | 0.019 | .885 | -0.063 | .671 |
| Age at blood draw, y | NA | 66 (5.2) | 0.011 | .420 | 0.015 | .448 | 0.070 | <.001 | 0.024 | .006 | 0.017 | <.001 | 0.058 | <.001 | 0.029 | <.001 | 0.034 | <.001 |
| BMI, kg/m² | NA | 27 (5.3) | -0.002 | .850 | -0.008 | .679 | 0.021 | .021 | 0.000 | .958 | 0.023 | <.001 | -0.040 | <.001 | 0.049 | <.001 | 0.028 | <.001 |
| Nonwhite race | 72 (8.8) | NA | -0.529 | .025 | -0.117 | .724 | 0.009 | .955 | -0.352 | .014 | 0.232 | .004 | -0.562 | <.001 | 0.092 | .457 | 0.188 | .178 |
| Family history | 339 (41.4) | NA | -0.047 | .735 | 0.089 | .648 | 0.123 | .211 | 0.080 | .341 | -0.062 | .192 | 0.095 | .198 | -0.015 | .832 | -0.012 | .884 |
| Oral contraceptive use | 306 (37.4) | NA | -0.092 | .542 | -0.032 | .879 | -0.093 | .384 | 0.096 | .292 | -0.073 | .158 | 0.081 | .312 | 0.021 | .788 | 0.027 | .759 |
| Nulliparous | 55 (6.7) | NA | -0.008 | .977 | -0.259 | .498 | 0.067 | .726 | 0.039 | .812 | -0.088 | .346 | 0.001 | .994 | 0.106 | .458 | 0.039 | .811 |
| History of endometriosis | 38 (4.6) | NA | 0.460 | .150 | 0.778 | .084 | 0.095 | .676 | 0.291 | .133 | 0.057 | .606 | 0.042 | .803 | -0.092 | .585 | 0.165 | .385 |
| Current smoker | 62 (7.6) | NA | -0.262 | .332 | -0.417 | .273 | 1.187 | <.001 | 0.869 | <.001 | 0.235 | .012 | 0.213 | .140 | -0.292 | .040 | 0.418 | .009 |
| Prior hysterectomy | 218 (26.7) | NA | -0.468 | .039 | -0.477 | .135 | -0.233 | .148 | -0.076 | .578 | -0.011 | .888 | -0.266 | .029 | -0.070 | .556 | -0.021 | .875 |
| Current hormone therapy with intact uterus | 234 (28.6) | NA | 0.133 | .430 | -0.044 | .851 | -0.045 | .709 | 0.027 | .789 | -0.135 | .020 | -0.391 | <.001 | 0.170 | .055 | -0.240 | .017 |
| Current hormone therapy with prior hysterectomy | 130 (15.9) | NA | 0.078 | .766 | 0.082 | .825 | -0.018 | .922 | 0.187 | .241 | -0.296 | .001 | -0.121 | .391 | 0.452 | .001 | -0.239 | .127 |

* BMI = body mass index; Coeff = coefficient; HE4 = human epididymis protein 4; IGFBP2 = insulin-like growth factor binding protein 2; MMP7 = matrix metalloproteinase 7; NA = not applicable; PEB = parametric empirical Bayes rule. Coefficients and P values were obtained using ordinary least squares regression by regressing marker levels on all of the personal characteristics simultaneously. Marker levels were log transformed to promote normality and then centered and scaled so that the randomly selected control subjects and family history control subjects together have a mean of 0 and an SD of 1 on the transformed scale. Coefficients can therefore be interpreted as the change in the marker level in terms of SDs of the marker in healthy control subjects expected with a 1-unit change in the predictor variable. Each coefficient and P value represent the estimated effect of the personal characteristic on the marker level after adjusting for the effects of all other personal characteristics in the table. All statistical tests were two-sided.

† The number and percentage is given for categorical variables including increasing CA125 (PEB) (yes or no on the basis of a personal threshold yielding 96.9% specificity), no increasing CA125 (yes or no), nonwhite race (yes or no), family history (breast or ovarian cancer, yes or no), oral contraceptive use (≥1 year, yes or no), nulliparous (yes or no), history of endometriosis (yes or no), current smoker (yes or no), prior hysterectomy (yes or no), current hormone therapy with intact uterus (yes or no), and current hormone therapy with prior hysterectomy (yes or no). The mean and SD is reported for continuous variables including age at blood draw (years) and BMI (kg/m²).

‡ A single indicator variable was used to indicate diagnosis with ovarian cancer when estimating the effect of cancer on expected CA125 levels. For all other markers, patients were stratified by whether or not CA125 was elevated in the same blood sample using a threshold corresponding to 96.9% specificity obtained by the parametric empirical Bayes rule.

Table 2. Cancers detected and specificity of increasing CA125, TVU, and HE4 alone and in combination, by tumor type and stage on the basis of 84 patients and 516 control subjects in the Prostate, Lung, Colon, and Ovarian trial with CA125, HE4, and TVU imaging results available*

| Patient or control subject group | No. | Alone as first-line screen, No. (%) | | | As a second-line screen with CA125, No. (%)† | | |
|----------------------------------|-----|-------------------------------------|------------|------------|--|------------|------------|
| | | CA125 | TVU | HE4 | TVU | HE4 | TVU or HE4 |
| Type I‡ | | | | | | | |
| Stage I or II | 16 | 7 (43.8) | 9 (56.2) | 3 (18.8) | 4 (25.0) | 3 (18.8) | 5 (31.2) |
| Stage III, IV, or unknown | 31 | 13 (41.9) | 7 (22.6) | 11 (35.5) | 5 (16.1) | 9 (29.0) | 11 (35.5) |
| Type II§ | | | | | | | |
| Stage I or II | 5 | 1 (20.0) | 1 (20.0) | 2 (40.0) | 1 (20.0) | 1 (20.0) | 1 (20.0) |
| Stage III or IV | 32 | 18 (56.2) | 7 (21.9) | 14 (43.8) | 7 (21.9) | 14 (43.8) | 14 (43.8) |
| Overall sensitivity | 84 | 39 (46.4) | 24 (28.6) | 30 (35.7) | 17 (20.2) | 27 (32.1) | 31 (36.9) |
| Overall specificity | 516 | 500 (96.9) | 500 (96.9) | 500 (96.9) | 515 (99.8) | 514 (99.6) | 513 (99.4) |

* HE4 = human epididymis protein 4; PEB = parametric empirical Bayes rule; TVU = transvaginal ultrasound.

† Sensitivities reported here reflect the percent of patients with increasing CA125 (PEB) with 96.9% specificity as evaluated by the PEB who were also positive for the listed marker at the same level of specificity.

‡ Includes all low-grade (grades 1 and 2) tumors and all tumors of clear cell, endometrioid, and mucinous histology.

§ Includes grade 3 and 4 tumors with serous, undifferentiated, or “adenocarcinoma not otherwise specified” histology.

of a morphology index (16) for TVU interpretation would likely improve overall screening performance.

Our analysis provides empiric support for the multimodal screening strategy that is being tested in the UK Collaborative Trial of Ovarian Cancer Screening. We conclude that HE4 may have a role in multimodal screening. However, measuring HE4 will not make screening itself effective. Requiring two tests both to be positive improves specificity, but use of an “and” rule can be expected to reduce sensitivity relative to use of an “or” rule. The efficacy of multimodal screening for ovarian cancer has not been demonstrated, and diagnosing HE4-associated tumors is not necessarily better than diagnosing TVU-associated tumors in terms of health outcomes, as there are no reports of one being more beneficial than the other in terms of outcome. Accordingly these results have more utility for future research than for clinical practice. Confirmation by an independent group that measurement of HE4 outperforms TVU as both a first- and second-line screen is needed.

References

- Partridge E, Kreimer AR, Greenlee RT, et al. Results from four rounds of ovarian cancer screening in a randomized trial. *Obstet Gynecol.* 2009;113(4):775–782.
- Buys SS, Partridge E, Black A, 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.
- 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(4):327–340.
- Palmer C, Duan X, Hawley S, et al. Systematic evaluation of candidate blood markers for detecting ovarian cancer. *PLoS One.* 2008;3(7):e2633.
- Simon I, Liu Y, Krall KL, et al. Evaluation of the novel serum markers B7-H4, Spondin 2, and DcR3 for diagnosis and early detection of ovarian cancer. *Gynecol Oncol.* 2007;106(1):112–118.
- Havrilesky LJ, Whitehead CM, Rubatt JM. Evaluation of biomarker panels for early stage ovarian cancer detection and monitoring for disease recurrence. *Gynecol Oncol.* 2008;110(3):374–382.
- Lancaster JM, Dressman HK, Clarke JP, et al. Identification of genes associated with ovarian cancer metastasis using microarray expression analysis. *Int J Gynecol Cancer.* 2006;16(5):1733–1745.
- Anderson GL, McIntosh MW, Wu L, et al. Assessing lead time of selected ovarian cancer biomarkers: a nested case-control study. *J Natl Cancer Inst.* 2010;102(1):26–38.
- Cramer DW, Bast RC Jr, Berg CD. Ovarian cancer biomarker performance in Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial specimens. *Cancer Prev Res (Phila).* 2011;4(3):365–374.
- Zhu CS, Pinsky PF, Cramer DW, et al.; for the PLCO Project Team. A framework for evaluating biomarkers for early detection: validation of biomarker panels for ovarian cancer. *Cancer Prev Res (Phila).* 2011;4(3):375–383.

- McIntosh M, Urban N, Karlan B. Generating longitudinal screening algorithms using novel biomarkers for disease. *Cancer Epidemiol Biomarkers Prev.* 2002;11(2):159–166.
- McIntosh M, Urban N. A parametric empirical Bayes method for cancer screening using longitudinal observations of a biomarker. *Biostatistics.* 2003;4(1):27–40.
- Lowe KA, Shah C, Wallace E, 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(9):2480–2487.
- Johnson CC, Kessel B, Riley TL, et al. The epidemiology of CA-125 in women without evidence of ovarian cancer in the Prostate, Lung, Colorectal and Ovarian Cancer (PLCO) Screening Trial. *Gynecol Oncol.* 2008;110(3):383–389.
- Kurman RJ, Shih IeM. The origin and pathogenesis of epithelial ovarian cancer: a proposed unifying theory. *Am J Surg Pathol.* 2010;34(3):433–443.
- DePriest PD, Varner E, Powell J, et al. The efficacy of a sonographic morphology index in identifying ovarian cancer: a multi-institutional investigation. *Gynecol Oncol.* 1994;55(2):174–178.

Funding

This work was supported by the Pacific Ovarian Cancer Research Consortium from the National Cancer Institute (P50 CA083636 to NU). Support for the Translational and Outcomes Research Laboratory from National Cancer Institute (U01 CA152637 to NU; Chris Li as primary investigator) and the Listwin Family and Canary Foundations (to NU) is also gratefully acknowledged.

Notes

The content is solely the responsibility of the authors and does not necessarily represent the official views

of the National Cancer Institute or the National Institute of Health.

Affiliations of authors: Molecular Diagnostics Program, Fred Hutchinson Cancer Research Center, Seattle, WA (NU, JDT, LAB, RMF, AVK, KCO'B, CWD); Computational Biology Program, Fred Hutchinson Cancer Research Center,

Seattle, WA (MWM); Cancer Prevention Program, Fred Hutchinson Cancer Research Center, Seattle, WA (GLA); University of Pennsylvania Perelman School of Medicine, Penn Ovarian Cancer Research Center, Philadelphia, PA (NS); Department of Obstetrics and Gynecology, Epidemiology Center, Brigham

and Women's Hospital, Boston, MA (DWC); Early Detection Research Group, Division of Cancer Prevention, National Cancer Institute, National Institutes of Health, Bethesda, MD (CDB); Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD (PH).