Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2016 September ; 25(9): 1333–1340. doi:
10.1158/1055-9965.EPI-15-1299.

Validation of a novel biomarker panel for the detection of ovarian cancer

Felix Leung^{1,2}, Marcus Q. Bernardini³, Marshall D. Brown⁴, Yingye Zheng⁴, Rafael Molina⁵, Robert C. Bast Jr.⁶, Gerard Davis⁷, Stefano Serra⁸, Eleftherios P. Diamandis^{1,2,9}, and Vathany Kulasingam^{1,9,*}

¹Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada

²Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ontario, Canada

³Division of Gynecologic Oncology, Department of Obstetrics and Gynecology, University of Toronto, Toronto, Ontario

⁴Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America

⁵Service of Clinical Biochemistry, Hospital Clínic de Barcelona, Barcelona, Spain

⁶Department of Experimental Therapeutics, University of Texas MD Anderson Cancer Center, Houston, Texas, United States of America

⁷Abbott Diagnostics, Lake Forest, Illinois, United States of America

⁸Department of Pathology, Laboratory Medicine Program, University Health Network, Toronto, Ontario, Canada

⁹Department of Clinical Biochemistry, University Health Network, Toronto, Ontario, Canada

Abstract

Background—Ovarian cancer (OvCa) is the most lethal gynecological malignancy. Our integrated -omics approach to OvCa biomarker discovery has identified kallikrein 6 (KLK6) and folate-receptor 1 (FOLR1) as promising candidates but these markers require further validation.

Methods—KLK6, FOLR1 CA125 and HE4 were investigated in three independent serum cohorts with a total of 20 healthy controls, 150 benign controls and 216 OvCa patients. The serum biomarker levels were determined by ELISA or automated immunoassay.

Results—All biomarkers demonstrated elevations in the sera of OvCa patients compared to controls ($p < 0.01$). Overall, CA125 and HE4 displayed the strongest ability (AUC 0.80 and 0.82, respectively) to identify OvCa patients and the addition of HE4 to CA125 improved the sensitivity from 36% to 67% at a set specificity of 95%. As well, the combination of HE4 and FOLR1 was a

*Corresponding author: Vathany Kulasingam, PhD, FCACB, Clinical Biochemist, Department of Clinical Biochemistry, Assistant Professor, University of Toronto, University Health Network, Laboratory Medicine Program, 200 Elizabeth Street, Room 3 EB 362A, Toronto, ON M5G 2C4, Phone: 416-340-4800 ext. 8589, Fax: 416-340-4215, vathany.kulasingam@uhn.ca.

Disclosures: Dr. Robert C. Bast, Jr. receives royalties for the discovery of CA125 from Fujirebio Diagnostics, Inc.

strong predictor of OvCa diagnosis, displaying comparable sensitivity (65%) to the best performing CA125-based models (67%) at a set specificity of 95%.

Conclusions—The markers identified through our integrated –omics approach performed similarly to the clinically-approved markers CA125 and HE4. Furthermore, HE4 represents a powerful diagnostic marker for OvCa and should be used more routinely in a clinical setting.

Impact—The implications of our study are two-fold: (1) we have demonstrated the strengths of HE4 alone and in combination with CA125, lending credence to increasing its usage in the clinic; and (2) we have demonstrated the clinical utility of our integrated –omics approach to identifying novel serum markers with comparable performance to clinical markers.

Keywords

ovarian cancer; biomarker; ELISA; early detection; diagnosis

INTRODUCTION

Ovarian cancer is the most lethal gynecological malignancy and the fifth-leading cause of mortality due to cancer in North American women. While the 5-year survival rate for cases diagnosed at an early stage (I–II) is approximately 80–90%, this decreases to 20–30% in late stage diagnoses (III–IV) [1]. Unfortunately, very few ovarian cancer cases are diagnosed at early stages while the tumor is still localized or confined to the ovary [2].

Since its discovery in 1981 by Bast, Jr. *et al.* [3], carbohydrate antigen 125 (CA125) still remains the gold-standard serum biomarker for ovarian cancer. CA125 is approved for both monitoring treatment with chemotherapy and differential diagnosis of patients presenting with a pelvic mass. The standard clinical cut-off value for CA125 is 35 U/mL, although serum levels have been shown to fluctuate depending on race, menstrual cycle timepoint, and presence of non-ovarian cancer pathologies [4–7]. As such, a major limitation of CA125 is that it displays poor specificity for ovarian cancer overall [8–10]. Additionally, CA125 is often not elevated in early-stage disease or in select subtypes of ovarian carcinoma such as mucinous neoplasms [11]. For these reasons, CA125 is not approved for ovarian cancer screening or for the detection of early disease on its own. The Prostate, Lung, Colorectal, and Ovarian (PLCO) and the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) screening trials represent two of the largest prospective trials worldwide that examined the clinical utility of CA125 in screening for ovarian cancer in asymptomatic women [12, 13]. The main objective of these trials were to demonstrate whether or not there is an overall survival benefit to screening asymptomatic women with ultrasound or with ultrasound plus CA125 versus no screening. Results for the PLCO trial have demonstrated that screening with CA125 and transvaginal ultrasound does not reduce mortality rates compared with standard care [14]. Meanwhile, the UKCTOCS trial randomly assigned approximately 200,000 post-menopausal women in a 1:1:2 ratio to annual multimodal screening (MMS) with serum CA125 interpreted with the risk of ovarian cancer algorithm (ROCA) and with transvaginal ultrasound (USS); annual USS alone; or no screening [15]. The study was powered to detect a mortality reduction of 30%. The primary outcome analysis spanning 0–14 years showed no significant reduction in mortality in the

MMS and USS groups (15% vs 11%) when compared to the no screening arm. Nonetheless, a secondary sub-group analysis did show the benefit of screening in women between the latter half of the screening period (years 7–14), when prevalent cases were excluded (28% mortality reduction after 7 years of screening in the MMS group). The authors state that additional follow-up of the UKCTOCS cohort is necessary before “firm conclusions” can be reached on the efficacy and cost-effectiveness of ovarian cancer screening. As such, novel algorithms and biomarkers that enable accurate prediction of the presence of ovarian malignancy in women are still being sought.

Previously, we have reported the potential utility of an integrated approach to ovarian cancer biomarker discovery [16]. This in-house approach to biomarker discovery was developed as a means of translating mass spectrometry-based proteomics to clinically relevant and meaningful biomarkers. To accomplish this, we complemented proteomic analyses of the conditioned media of ovarian cancer cell lines [17] and ascites fluid [18] with transcriptomics and computational biology in order to capture the entirety of the disease and extract the most promising candidates for serum validation. To this end, we have successfully validated one of the putative markers identified through this integrated approach. We reported significant elevations of folate receptor 1 (FOLR1) in the serum of ovarian cancer patients compared to healthy controls and patients with benign gynaecological conditions in a preliminary validation cohort [19]. The successful validation of FOLR1 served as a proof-of-principle of our integrated approach to identifying novel ovarian cancer biomarkers. Following this scheme, kallikrein 6 (KLK6) was also identified as a putative serum marker for ovarian cancer with diagnostic utility similar to that of the FDA-approved markers CA125 and HE4 (data not shown).

In this study, we investigated the levels of KLK6 and FOLR1 along with the FDA-approved markers, CA125 and HE4, in three independent serum cohorts consisting of a spectrum of ovarian cancer patients and patients with non-ovarian cancer pathologies. The main objective was to determine the ability of our putative markers to differentiate ovarian cancer from non-ovarian cancer patients and to identify any additional clinically-meaningful stratifications using these novel markers in a multiparametric algorithm.

MATERIALS AND METHODS

Patients and specimens

The study population was comprised of three independent serum cohorts: (1) a Spanish retrospective cohort, (2) an American blinded retrospective cohort, and (3) a Canadian blinded prospective cohort. To retrieve serum from patients, blood was collected in BD vacutainer tubes (SST) and allowed to clot for 30 minutes at room temperature. Samples were then centrifuged at room temperature for 10 minutes at 1500 g to pellet the cells. Immediately following centrifugation, the sera were aliquoted into 1 mL cryotubes and stored at –80 °C until analysis. All samples were collected with informed consent and Institutional Review Board approval.

The Spanish retrospective cohort consisted of 100 serum samples collected in Barcelona, Spain according to standardized protocols mentioned above. The samples were collected

pre-operatively and prior to treatment from women with gynaecological disease, with 50 serum samples from patients later identified as having been diagnosed with a spectrum of non-ovarian cancer pathologies and 50 serum samples from patients with histologically diagnosed ovarian cancer.

The American blinded retrospective cohort consisting of 60 samples was collected in Houston, USA according to standardized protocols mentioned above. The healthy control samples were collected as described in a previous study [20]. The remaining samples were collected preoperatively after imaging of patients with benign disease and ovarian cancer. A total of 60 serum samples were collected from 20 patients established as healthy controls, 20 patients diagnosed with non-ovarian cancer pathologies, and 20 patients diagnosed with ovarian cancer.

The Canadian blinded prospective cohort consisting of 226 samples was collected in Toronto, Canada according to standardized protocols mentioned above. The samples were collected according to the prospective-specimen-collection, retrospective-blinded-evaluation guidelines outlined by Pepe *et al* [21] from a population of high-risk patients referred to a tertiary care centre. A total of 80 patients were eventually diagnosed with non-ovarian cancer pathologies while a total of 146 patients were eventually diagnosed with histologically-confirmed ovarian cancer.

No patients with family history of cancer and/or positive for *BRCA1/BRCA2* mutations were included in the three cohorts. Specific details on the ethnicities of patients were unavailable but it can be extrapolated that the American and Canadian cohorts were reflective of the populations and contained a mixture of ethnicities. Ovarian cancer staging was determined according to the FIGO classification system [22]. Ovarian cancer grading was determined based on cellular appearance: Grade 1 denoted well-differentiated cells, Grade 2 denoted moderately differentiated cells, and Grade 3 denoted poorly differentiated cells. The details of the three validation cohorts are outlined in Table 1.

Measurement of serum biomarkers

CA125 and HE4 were measured with commercially-available clinical grade immunoassays on the Abbott Architect i2000 platform (Abbott Diagnostics). The precision of these assays was <5–7%.

KLK6 was measured using an in-house developed sandwich enzyme-linked immunosorbent assay (ELISA) as described previously [23]. Briefly, KLK6-specific monoclonal antibody (clone 27–4) was immobilized in 96-well microtiter plates by incubating 500 ng/well in coating buffer (50 mM Tris, pH 7.8) overnight. After washing with washing buffer (5 mM Tris, 150 mM NaCl, 0.05% Tween-20, pH 7.8), 50 µL of serum or standards was pipetted into each well and incubated with 50 µL of assay buffer (6% BSA, 50 mM Tris, 10% goat IgG, 2% mouse IgG, 1% bovine IgG, 0.5 M KCl, pH 7.8) for 1 hour. Following washing, another biotinylated KLK6-specific monoclonal antibody (code E24) was added (50 ng/well) and incubated for 1 hour. Unbound biotinylated antibody was washed off and 100 µL of streptavidin-conjugated horseradish peroxidase was then added for 15 minutes. A final washing preceded the addition of 100 µL of 3,3',5,5'-tetramethylbenzidine (TMB) substrate

solution to develop the quantifiable signal and the chromogenic reaction was stopped with the addition of stop solution (2 N H₂SO₄) after 10 minutes of incubation. Absorbance was measured with the Wallac EnVision 2103 Multilabel Reader (Perkin Elmer) at 450 nm standardized to background absorbance at 620 nm. All serum samples were diluted 5-fold with 60 g/L bovine serum albumin before analysis.

FOLR1 was measured using a commercially-available sandwich ELISA kit (R&D Systems). Briefly, FOLR1-specific mouse anti-human capture antibody was immobilized in 96-well microtiter plates by incubating 200 ng/well in 1X PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.4) overnight. After washing with washing buffer, the plates were loaded with 10 g/L bovine serum albumin in 1X PBS for 1 hour to block any non-specific interactions. The plates were then aspirated and loaded with 100 µL of serum or standards and incubated for 2 hours. Following washing, goat anti-human biotinylated detection antibody was added (90 ng/well) and incubated for 1 hour. Unbound biotinylated antibody was washed off and 100 µL of streptavidin-conjugated horseradish peroxidase was then added for 15 minutes. Similar to the KLK6 assay, a final washing preceded the addition of 100 µL of TMB substrate solution and the chromogenic reaction was stopped with the addition of stop solution after 10 minutes of incubation. Absorbance was measured at 450 nm standardized to background absorbance at 620 nm. All serum samples were diluted 5-fold with 10 g/L bovine serum albumin before analysis.

Statistical analysis

Medians and interquartile ranges were used to report age and biomarker values in normal, benign, and cancerous serum from subjects in each cohort and marginally across cohorts. The associations between biomarker values and outcome were evaluated with logistic regression models, and log odds ratios with corresponding 95% confidence intervals (CI) in each cohort and marginally across cohorts were reported. For models where data were combined across cohorts, a fixed cohort effect was included. All multivariate models were age-adjusted. Biomarkers were log transformed in all model-fitting procedures.

Diagnostic performance for each biomarker and in combination was evaluated using receiver characteristic operating (ROC) curve analysis, and by reporting the area-under-the-curve (AUC) and by examining sensitivity at fixed specificity levels. Again, performance was estimated in each cohort and marginally across cohorts. For multivariate models, performance summaries were calculated based on the linear predictor including biomarker values and age. Since multivariate models were fit on the same data used for diagnostic evaluation, we employed 5-fold cross-validation to estimate all performance metrics. Confidence intervals were calculated using 1,000 bootstrap replicates accounting for variations in both model fitting and estimating accuracy measures. The performance of linear scores from multivariate models from all data were also evaluated on pre-specified subgroups.

RESULTS

Diagnostic performances of biomarkers

In all three validation cohorts, our in-house markers (FOLR1 and KLK6) and the FDA-approved markers (CA125 and HE4) all displayed elevations in the serum of ovarian cancer patients compared to patients with non-ovarian cancer pathologies (Figure 1). These elevations were all found to be significant ($p < 0.01$) using nonparametric Mann-Whitney U and Kruskal-Wallis tests. The associations between biomarker values and the diagnosis of ovarian cancer were further evaluated using log odds ratios (Supplementary Table 1). For all biomarkers, significant associations ($p < 0.01$) were observed with ovarian cancer within each cohort and in all cohorts combined. When assessing the diagnostic ability of the markers using ROC curve analysis, HE4 and CA125 displayed the strongest ability to discriminate ovarian cancer patients from patients with non-ovarian pathologies, with AUCs of 0.82 and 0.80, respectively, in all cohorts combined (Figure 2). FOLR1 showed an AUC of 0.76 and KLK6 had an AUC of 0.56.

Interestingly, the AUC performances of the markers varied across the three independent cohorts. As seen in Table 2, CA125 and HE4 were significantly stronger predictors of ovarian cancer in the Spanish cohort compared to in the American cohort and Canadian cohort. FOLR1 performed similarly in the Spanish cohort as the Canadian cohort but displayed comparable diagnostic ability to CA125 and HE4 in the American cohort. A similar pattern was observed for KLK6, which performed similarly in the Spanish cohort as the Canadian cohort but in the American cohort, displayed comparable diagnostic ability as CA125 and HE4. These variations could have been due to the differences in the characteristics between the cohorts: whereas the Spanish and American cohorts were comprised of selected retrospective samples, the Canadian cohort (larger sample number) was prospectively collected and evaluated in a blinded fashion to case-control status. Furthermore, the Spanish and American samples originated from primarily late-stage patients diagnosed with serous carcinoma while the Canadian cohort was comprised of a more diverse population with respect to cancer histotype and staging (Table 1). As such, the Canadian cohort may provide the most accurate representation of how the biomarkers perform in the general population.

Although CA125, HE4 and FOLR1 displayed relatively similar performances as measured by the AUC, inspection of the univariate sensitivities at set specificities revealed important differences between the biomarkers. From visual inspection of univariate ROC curves (Figure 2), it is notable that for specificities ranging from 0.75 to 1, HE4 and FOLR1 have higher corresponding sensitivities compared to CA125. CA125 reaches highest sensitivities only when specificities are below 0.75. As seen in Table 2, HE4 displayed the highest sensitivity (57%) followed by FOLR1 (47%), CA125 (38%) and finally KLK6 (19%) at a set specificity of 95% in all cohorts combined. These observations held true even at a set specificity of 99% with HE4 displaying a sensitivity of 35%, followed by FOLR1 with 30%, CA125 with 20% and KLK6 with 14%. In the Spanish and Canadian cohorts, similar trends were observed. At a set specificity of 95%, HE4 displayed the highest sensitivities followed by FOLR1, CA125 and KLK6 (Supplementary Table 2). As well, CA125 and HE4 displayed

higher sensitivities in the Spanish cohort compared to the American cohort and the Canadian cohort at a fixed specificity of 95%.

Multivariate modeling for the diagnosis of ovarian cancer

Using a baseline model of CA125, we used multivariate modeling to assess the ability of KLK6, FOLR1 and HE4 to improve upon the gold-standard marker. As seen in Supplementary Table 3, the addition of either HE4 or FOLR1 was able to improve upon CA125-based diagnosis of ovarian cancer (both $p < 0.001$). The addition of KLK6 to CA125, however, did not add any significant value to CA125. In a multivariate model combining all four markers, it was observed that the addition of FOLR1 and KLK6 did not add any value to the combination CA125 and HE4. Interestingly, it was found that the combination of HE4 and FOLR1 demonstrated a significant association ($p < 0.001$) with the diagnosis of ovarian cancer when investigating multivariate models that did not incorporate CA125.

Overall, ROC curve analysis confirmed observations from the multivariate model odds ratios. As seen in Table 3, the addition of HE4, FOLR1 and KLK6 did not significantly improve the performance of CA125 in all cohorts combined. A model combining all four markers displayed equal diagnostic ability (AUC 0.85) to the combination of only CA125 and HE4 (AUC 0.85), indicating the lack of value FOLR1 and KLK6 adds to the combination of CA125 and HE4. However, the combination of HE4 and FOLR1 displayed comparable diagnostic ability (AUC 0.83) to the CA125-based models. Within each cohort, the CA125-based models displayed the strongest performances in the Spanish cohort (AUCs 0.91–0.94), followed by the American cohort (AUCs 0.79–0.86) and finally, displayed the lowest AUCs (0.73–0.79) in the Canadian cohort. Again, these trends may have been due to differences in patient composition within each cohort.

Further evaluation of the sensitivities at stringent specificities revealed specific patterns (Table 3). At a set specificity of 95%, the addition of HE4 to CA125 was able to yield the highest sensitivity of 67% compared to only 36% using CA125. The addition of KLK6 produced negligible differences (sensitivity of 38%) but the addition of FOLR1 resulted in a significant improvement in the sensitivity to 57% at a set specificity of 95%. The combination of HE4 and FOLR1 displayed a sensitivity of 65%, which was comparable with the best-performing model of CA125 and HE4 (67%) at a set specificity of 95%. These observations held true at a set specificity of 99% where HE4 and FOLR1, with a sensitivity of 39%, performed similarly to the strongest CA125-based model (all four markers combined) with a sensitivity of 38%.

When evaluating the performances of multivariate models with pre-defined cut-offs, both CA125 (cut-off defined as 35 U/mL) and HE4 (cut-off defined as 70 pmol/L) performed comparably in terms of the true positive fraction (TPF); CA125 displayed a TPF of 85% while HE4 displayed a TPF of 81% as seen in Table 4. However, the strength of HE4 was apparent when assessing the false positive fraction (FPF) as HE4 displayed a markedly lower FPF than CA125. This highlights the improved specificity HE4 displayed compared to CA125 without significant compromise to sensitivity. Finally, when HE4 was added to a model including CA125 and age there was a >2-fold decrease in FPF with only a slight

reduction in TPF. The addition of KLK6 or FOLR1 to CA125- and HE4-based models did not result in any significant improvements to the TPF and FPF.

Association of biomarkers with clinical characteristics

Univariate modeling revealed that all four markers in all cohorts combined demonstrated comparable AUC for early-stage diagnosis (tumor stage I/II) of approximately 0.65 (Supplementary Table 4). CA125, FOLR1 and HE4 correlated with tumor burden as shown by the fact that AUCs increased to 0.83–0.92 for tumor stage III/IV. Multivariate modeling improved the AUCs for early-stage diagnosis to approximately 0.75, although none of the models were able to outperform the baseline model of CA125. Similar to univariate modeling, multivariate modeling for late-stage disease showed improved AUC overall with the combination of CA125 with HE4 having the best performance with an AUC of 0.90.

In addition to tumor burden, all markers appeared to be more specific to the serous histotype. As seen in Supplementary Table 4, CA125 and HE4 showed the highest AUCs for the non-serous subtypes (0.71 and 0.73, respectively). When stratifying for only serous patients, the AUCs for all four markers improved significantly. Various multivariate models showed similar AUCs for the diagnosis of both non-serous (AUCs 0.74–0.79) and serous patients (AUCs 0.87–0.90).

DISCUSSION

In this study, we were able to demonstrate that the in-house markers identified through our integrated –omics approach displayed relatively strong and, in some instances, comparable performances to the FDA-approved markers for the discrimination of ovarian cancer patients. As seen by the consistent elevations of KLK6 and FOLR1 through the three independent cohorts, we have demonstrated the utility of our approach to integrating our in-house proteomic candidates with external high-throughput databases and subsequently identifying candidates with high potential to be novel serum biomarkers. Our in-house marker FOLR1 displayed diagnostic ability comparable to that of the FDA-approved marker CA125. Since our preliminary validation of FOLR1 as a serum marker for ovarian cancer, other groups have confirmed the diagnostic value of this marker. In study investigating serum folate receptor alpha (also known as FOLR1), mesothelin, and megakaryocyte potentiating factor, FOLR1 displayed a similar AUC (0.77) to that of CA125 (0.84) and HE4 (0.77) with respect to identifying ovarian cancer patients from healthy controls. Logistic regression modeling also demonstrated some complementarity between FOLR1 and CA125, in which a model combining the two markers with mesothelin and megakaryocyte potentiating factor displayed an AUC of 0.91 [24]. In a similar study by Kurosaki *et al.*, not only did serum FOLR1 demonstrate comparable diagnostic ability to CA125, but as well, serum FOLR1 consistently demonstrated higher specificity than the FDA-approved marker without significant compromise to sensitivity [25]. These findings are relatively consistent with our observations of FOLR1 in this study.

Another important conclusion from our study is the fact that HE4 consistently performed as the strongest marker in identifying ovarian cancer patients and showed the ability to improve upon the gold-standard marker CA125. HE4 displayed an overall AUC of 0.82 which was

comparable to that of CA125 (AUC 0.80) and displayed higher sensitivities than CA125 at stringent specificities. This held true for a multivariate model of CA125 and HE4 that displayed a higher sensitivity than CA125 alone. The combination of HE4 and CA125 was also the strongest multivariate model, outperforming the model combining all four markers. Clearly, HE4 is a powerful marker that also has the ability to improve CA125-based diagnosis of ovarian cancer patients. Despite the numerous studies investigating HE4 as an ovarian cancer marker and FDA-approval for the detection of cancer recurrence, this marker has not seen substantial penetrance in clinical use [26–28]. Our results in the three validation cohorts lend credence to this observation that numerous studies have noted – that HE4 performs strongly as a marker for ovarian cancer diagnosis and increased clinical use of this serum marker could potentially improve management of ovarian cancer.

As mentioned previously, the UKCTOCS trial found that the use of a multimodal strategy using longitudinal serum CA125 values resulted in a mortality reduction of 20%. The inspection of CA125 values over a period of time is based on the Risk of Ovarian Cancer Algorithm (ROCA) and was developed as an alternative to the limitations observed with using a fixed CA125 cut-off value [20]. Although a relatively novel concept, the ROCA has already been observed to be a potentially efficacious screening modality. In a single-arm prospective study of postmenopausal women, a 2-stage screening strategy that used ROCA to triage women for further transvaginal ultrasound displayed a positive predictive value of 40% at a specificity of 99.9%. Combined with the results of the UKCTOCS trial, the ROCA may represent a novel screening test for ovarian cancer and as well, a paradigm shift in the inspection of biomarkers. However, further evaluation of ROCA is required as there still remains unaddressed issues, such as the fact that no study has demonstrated the improvement in mortality reduction when comparing ROCA to using a fixed CA125 cutoff value [29].

There is now an increased focus on identifying multiparametric tests for the management of ovarian cancer rather than the identification of single biomarkers that will exceed the performance of CA125. The serum-based Risk of Malignancy Algorithm (ROMA) and OVA1™ tests also represent relatively new multimarker algorithms for the management of ovarian cancer. Although both tests have seen some success leading up to FDA-approval, there are still conflicting results as to how well they truly perform compared to CA125 [30–34]. Furthermore, due to the fact that the algorithms rely heavily on serum levels of CA125 and/or HE4, they are susceptible to bias towards patients with late-stage cancers and/or serous carcinoma (most often high-grade serous carcinoma) [35]. The preference of CA125 and HE4 for serous carcinoma patients is evident in our results as the two markers performed much better in cohorts with a greater proportion of samples originating from serous carcinoma patients. To complicate this matter, serous carcinoma rarely presents during early stages and so by the time pelvic mass patients are diagnosed through these algorithms, the cancer is (most likely) in an advanced stage.

Future studies will thus need to address the identification of biomarkers for patients diagnosed with early stage (I/II) ovarian cancer and patients diagnosed with non-serous ovarian carcinomas. With the appropriate discovery specimens, our integrated –omics approach can be applied in a similar manner to identify biomarkers specific for mucinous,

endometrioid and clear cell ovarian carcinoma. Not only are biomarkers for the detection of non-serous tumours lacking but also, non-serous tumours have a much greater chance of being detected during the early stages compared to their high-grade serous counterparts due to their indolent nature. The introduction of such subtype-specific biomarkers could markedly increase the survival of these underrepresented patients, as the current clinical markers are unable to detect these tumours with sufficient sensitivity and specificity.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial support: This work is supported by Proteomic Methods Inc., Toronto, Canada. Eleftherios P. Diamandis is funded by the Early Detection Research Network of NIH (EDRN; Grant #5U01CA152755). Yingye Zheng is supported by R01-GM085047 of NIH.

We thank Dr. Gerard Davis from Abbott Diagnostics for providing the CA125 and HE4 kits.

References

1. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. *CA Cancer J Clin.* 2010; 60:277–300. [PubMed: 20610543]
2. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin.* 2013; 63:11–30. [PubMed: 23335087]
3. Bast RC Jr, Feeney M, Lazarus H, Nadler LM, Colvin RB, Knapp RC. Reactivity of a monoclonal antibody with human ovarian carcinoma. *J Clin Invest.* 1981; 68:1331–7. [PubMed: 7028788]
4. Bast RC Jr, Xu FJ, Yu YH, Barnhill S, Zhang Z, Mills GB. CA 125: the past and the future. *Int J Biol Markers.* 1998; 13:179–87. [PubMed: 10228898]
5. Jacobs I, Bast RC Jr. The CA 125 tumour-associated antigen: a review of the literature. *Hum Reprod.* 1989; 4:1–12. [PubMed: 2651469]
6. Davelaar EM, van Kamp GJ, Verstraeten RA, Kenemans P. Comparison of seven immunoassays for the quantification of CA 125 antigen in serum. *Clin Chem.* 1998; 44:1417–22. [PubMed: 9665418]
7. Diamandis EP, Fritsche H, Lilja H, Chan DW, Schwartz MK. Tumor Markers: Physiology, Pathobiology, Technology, and Clinical Applications. American Association for Clinical Chemistry; 2002.
8. Sari R, Yildirim B, Sevinc A, Hilmioğlu F, Re. Zuckerman et al. --Sensitivity of CA-125 in patients with liver cirrhosis in the presence of ascites. *Am J Gastroenterol.* 2001; 96:253–4. [PubMed: 11197269]
9. Sevinc A, Adli M, Kalender ME, Camci C. Benign causes of increased serum CA-125 concentration. *Lancet Oncol.* 2007; 8:1054–5. [PubMed: 18054877]
10. Sevinc A, Buyukberber S, Sari R, Kiroglu Y, Turk HM, Ates M. Elevated serum CA-125 levels in hemodialysis patients with peritoneal, pleural, or pericardial fluids. *Gynecol Oncol.* 2000; 77:254–7. [PubMed: 10785474]
11. Leung F, Diamandis EP, Kulasingam V. From bench to bedside: discovery of ovarian cancer biomarkers using high-throughput technologies in the past decade. *Biomark Med.* 2012; 6:613–25. [PubMed: 23075239]
12. Simpson NK, Johnson CC, Ogden SL, Gamito E, Trocky N, McGuire C, et al. Recruitment strategies in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial: the first six years. *Control Clin Trials.* 2000; 21:356S–78S. [PubMed: 11189688]
13. Menon U, Jacobs I. Screening for ovarian cancer. *Best Pract Res Clin Obstet Gynaecol.* 2002; 16:469–82. [PubMed: 12413929]

14. Buys SS, Partridge E, Black A, Johnson CC, Lamerato L, Isaacs C, et al. Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. *JAMA*. 2011; 305:2295–303. [PubMed: 21642681]
15. Jacobs IJ, Menon U, Ryan A, Gentry-Maharaj A, Burnell M, Kalsi JK, et al. Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet*. 2016; 387:945–56. [PubMed: 26707054]
16. Kulasingam V, Pavlou MP, Diamandis EP. Integrating high-throughput technologies in the quest for effective biomarkers for ovarian cancer. *Nat Rev Cancer*. 2010; 10:371–8. [PubMed: 20383179]
17. Gunawardana CG, Kuk C, Smith CR, Batruch I, Soosaipillai A, Diamandis EP. Comprehensive analysis of conditioned media from ovarian cancer cell lines identifies novel candidate markers of epithelial ovarian cancer. *J Proteome Res*. 2009; 8:4705–13. [PubMed: 19663500]
18. Kuk C, Kulasingam V, Gunawardana CG, Smith CR, Batruch I, Diamandis EP. Mining the ovarian cancer ascites proteome for potential ovarian cancer biomarkers. *Mol Cell Proteomics*. 2009; 8:661–9. [PubMed: 19047685]
19. Leung F, Dimitromanolakis A, Kobayashi H, Diamandis EP, Kulasingam V. Folate-receptor 1 (FOLR1) protein is elevated in the serum of ovarian cancer patients. *Clin Biochem*. 2013; 46:1462–8. [PubMed: 23528302]
20. Lu KH, Skates S, Hernandez MA, Bedi D, Bevers T, Leeds L, 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; 119:3454–61. [PubMed: 23983047]
21. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. *J Natl Cancer Inst*. 2008; 100:1432–8. [PubMed: 18840817]
22. Prat J. Staging classification for cancer of the ovary, fallopian tube, and peritoneum. *Int J Gynaecol Obstet*. 2014; 124:1–5. [PubMed: 24219974]
23. Diamandis EP, Yousef GM, Soosaipillai AR, Bunting P. Human kallikrein 6 (zyme/protease M/neurosin): a new serum biomarker of ovarian carcinoma. *Clin Biochem*. 2000; 33:579–83. [PubMed: 11124344]
24. O'Shannessy DJ, Somers EB, Palmer LM, Thiel RP, Oberoi P, Heath R, et al. Serum folate receptor alpha, mesothelin and megakaryocyte potentiating factor in ovarian cancer: association to disease stage and grade and comparison to CA125 and HE4. *J Ovarian Res*. 2013; 6:29. [PubMed: 23590973]
25. Kurosaki A, Hasegawa K, Kato T, Abe K, Hanaoka T, Miyara A, et al. Serum folate receptor alpha as a biomarker for ovarian cancer: Implications for diagnosis, prognosis and predicting its local tumor expression. *Int J Cancer*. 2016; 138:1994–2002. [PubMed: 26595060]
26. Granato T, Porpora MG, Longo F, Angeloni A, Manganaro L, Anastasi E. HE4 in the differential diagnosis of ovarian masses. *Clin Chim Acta*. 2015; 446:147–55. [PubMed: 25892674]
27. Michalak M, Gasiorowska E, Markwitz EN. Diagnostic value of CA125, HE4, ROMA and logistic regression model in pelvic mass diagnostics - our experience. *Ginekol Pol*. 2015; 86:256–61. [PubMed: 26117983]
28. Jia LT, Zhang YC, Li J, Tian Y, Li JF. The role of human epididymis protein 4 in the diagnosis of epithelial ovarian cancer. *Clin Transl Oncol*. 2016; 18:233–9. [PubMed: 26220095]
29. Narod SA, Sopik V, Giannakeas V. Should we screen for ovarian cancer? A commentary on the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) randomized trial. *Gynecol Oncol*. 2016; 141:191–4. [PubMed: 26915499]
30. Bast RC Jr, Skates S, Lokshin A, Moore RG. Differential diagnosis of a pelvic mass: improved algorithms and novel biomarkers. *Int J Gynecol Cancer*. 2012; 22(Suppl 1):S5–S8. [PubMed: 22543921]
31. Grenache DG, Heichman KA, Werner TL, Vucetic Z. Clinical performance of two multi-marker blood tests for predicting malignancy in women with an adnexal mass. *Clin Chim Acta*. 2015; 438:358–63. [PubMed: 25283731]

32. Lokich E, Palisoul M, Romano N, Craig MM, Robison K, Stuckey A, et al. Assessing the risk of ovarian malignancy algorithm for the conservative management of women with a pelvic mass. *Gynecol Oncol.* 2015; 139:248–52. [PubMed: 26364809]
33. Karlsen MA, Sandhu N, Hogdall C, Christensen IJ, Nedergaard L, Lundvall L, et al. Evaluation of HE4, CA125, risk of ovarian malignancy algorithm (ROMA) and risk of malignancy index (RMI) as diagnostic tools of epithelial ovarian cancer in patients with a pelvic mass. *Gynecol Oncol.* 2012; 127:379–83. [PubMed: 22835718]
34. Braicu EI, Van GT, Nassir M, Richter R, Chekerov R, Gasimli K, et al. Preoperative HE4 and ROMA values do not improve the CA125 diagnostic value for borderline tumors of the ovary (BOT) - a study of the TOC Consortium. *J Ovarian Res.* 2014; 7:49. [PubMed: 24872845]
35. Lennox GK, Eiriksson LR, Reade CJ, Leung F, Mojtahedi G, Atenafu EG, et al. Effectiveness of the risk of malignancy index and the risk of ovarian malignancy algorithm in a cohort of women with ovarian cancer: does histotype and stage matter? *Int J Gynecol Cancer.* 2015; 25:809–14. [PubMed: 25855958]

Figure 1.

Distributions of CA125, HE4, FOLR1 and KLK6 in the three validation cohorts. The serum marker levels are plotted against a logarithmic scale with the horizontal lines representing the medians. * denotes $P < 0.01$ by Mann-Whitney U and Kruskal-Wallis tests.

Figure 2.

Univariate ROC curve analysis of CA125, HE4, FOLR1 and KLK6 in the three validation cohorts combined. Serum marker levels were assessed for their ability to discriminate ovarian cancer from non-ovarian cancer patients. The vertical grey lines represent set specificities of 99% and 95% from left to right. The AUC values are shown to the right with the associated 95% confidence intervals.

Table 1

Clinical characteristics of the samples in the three validation cohorts.

Sample characteristics	Spanish cohort	American cohort	Canadian cohort	All cohorts
<i>Diagnosis</i>				
Healthy control	0	20	0	20
Benign disease	50	20	80	150
Ovarian cancer	50	20	146	216
<i>Histotype (Cancer)</i>				
Serous	35	12	78	125
Clear cell	4	3	11	18
Endometrioid	4	1	17	22
Mucinous	4	0	16	20
Other	3	4	24	31
<i>Histotype (Benign)</i>				
Serous neoplasm	17	6	24	47
Endometrioid/endometrial neoplasm	13	6	17	36
Mucinous neoplasm	0	1	9	10
Non-epithelial neoplasm	19	3	24	46
Brenner tumor	0	0	2	2
Other	1	4	4	9
<i>Staging</i>				
Borderline	3	0	0	3
I/II	4	1	65	70
III/IV	43	19	79	141
Unknown	0	0	2	2
<i>Grading</i>				
1	N/A	N/A	13	13
2	N/A	N/A	10	10
3	N/A	N/A	69	69
Unknown	N/A	N/A	38	38
<i>Total</i>	100	60	226	386
<i>Median age (Interquartile range)</i>	51 (41,66)	60 (48,65)	57 (48,66)	57 (46,66)

Table 2

Univariate analysis of CA125, HE4, FOLR1, and KLK6 for the diagnosis of ovarian cancer cases.

Marker	AUC (95% CI)			Sensitivity (95% CI) at set specificity of:	
	Spanish cohort	American cohort	Canadian cohort	All cohorts	
CA125	0.90 (0.83–0.95)	0.83 (0.69–0.95)	0.73 (0.66–0.79)	0.80 (0.76–0.85)	38.0% (24.2–49.5) 20.4% (15.2–31.4)
HE4	0.92 (0.85–0.97)	0.82 (0.68–0.93)	0.79 (0.74–0.85)	0.82 (0.78–0.86)	56.5% (41.1–67.5) 35.2% (30.5–50.2)
FOLR1	0.76 (0.64–0.86)	0.80 (0.64–0.94)	0.74 (0.67–0.80)	0.76 (0.71–0.81)	47.2% (39.8–62.1) 29.6% (22.5–50.2)
KLK6	0.68 (0.56–0.79)	0.80 (0.63–0.93)	0.63 (0.56–0.71)	0.56 (0.51–0.62)	19.0% (13.2–24.8) 14.4% (9.6–21.0)

Table 3

Multivariate analysis of CA125, HE4, FOLR1, and KLK6 for the diagnosis of ovarian cancer cases.

Marker	AUC (95% CI)				Sensitivity (95% CI) at set specificity of:		
	Spanish cohort	American cohort	Canadian cohort	All cohorts	95%	99%	
CA125	0.91 (0.84–0.96)	0.80 (0.66–0.93)	0.74 (0.67–0.80)	0.82 (0.78–0.86)	35.6% (19.4–49.0)	19.0% (10.2–31.5)	
CA125 + HE4	0.94 (0.88–0.98)	0.79 (0.66–0.91)	0.79 (0.74–0.85)	0.85 (0.81–0.89)	67.1% (38.4–58.8)	37.0% (9.7–56.9)	
CA125 + FOLR1	0.91 (0.84–0.97)	0.83 (0.69–0.93)	0.74 (0.69–0.81)	0.83 (0.79–0.87)	56.9% (34.7–63.9)	30.1% (14.9–46.4)	
CA125 + KLK6	0.91 (0.83–0.97)	0.83 (0.69–0.94)	0.73 (0.66–0.80)	0.82 (0.78–0.86)	38.4% (24.9–52.3)	25.5% (16.2–39.3)	
All	0.92 (0.86–0.98)	0.86 (0.74–0.97)	0.79 (0.74–0.85)	0.85 (0.82–0.89)	51.9% (43.1–63.9)	38.4% (16.6–62.0)	
HE4 + FOLR1	0.89 (0.83–0.96)	0.84 (0.71–0.93)	0.79 (0.73–0.85)	0.83 (0.80–0.87)	64.8% (40.3–61.6)	38.8% (17.5–63.8)	

All models are age-adjusted

Table 4

Performances of multivariate models in all cohorts combined based on defined thresholds for biomarkers. Cut-off values for KLK6 and FOLR1 were calculated using the 75th percentile marker value in non-cancerous tissues. 95% confidence intervals were calculated using the bootstrap.

Rule	Sensitivity or TPF* (95% CI)	100%-Specificity or FPF* (95% CI)
CA125>35¹	85% (80–89)	40% (33–47)
+ HE4>70 ³	75% (69–80)	15% (10–21)
+ FOLR1>0.775 ²	63% (56–69)	8% (4–12)
+ KLK6>2.075 ²	37% (31–44)	6% (3–11)
+ KLK6>2.075 ² , FOLR1>0.775 ² and HE4>70 ³	31% (26–38)	2% (0–4)
HE4>70³	81% (75–86)	32% (25–39)
+ FOLR1>0.775 ²	62% (56–68)	12% (7–18)
+ KLK6>2.075 ²	36% (30–42)	9% (5–14)
+ CA125>35 ¹ and FOLR1>0.775 ²	60% (53–66)	5% (2–9)

* TPF = true positive fraction;

° FPF = false positive fraction

¹ U/mL;

² ng/mL;

³ pmol/L