Detection of the HE4 Protein in Urine as a Biomarker for Ovarian Neoplasms

Ingegerd Hellstrom1,*, Patrick J. Heagerty2, Elizabeth M. Swisher3, Pu Liu1, Jade Jaffar1, Kathy Agnew3, and Karl Erik Hellstrom1

1Department of Pathology, University of Washington, Seattle, WA 98104, USA.
2Department of Biostatistics, University of Washington, Seattle, WA 98104, USA.
3Department of Obstetrics/Gynecology, University of Washington, Seattle, WA 98104, USA.

Abstract

The HE4 protein is overexpressed in ovarian carcinomas and can be detected in serum by an ELISA with sensitivity similar to CA125 and higher specificity for malignant disease. We now demonstrate that HE4 can also be detected in the urine at a specificity level of 94.4%, including 13/15 (86.6%) with stage I/II and 57/64 (89.0%) with stage III/IV disease and including 90.5% of patients with serous ovarian carcinoma. Assaying serum and urine from the same patients showed similar sensitivity. Our data indicate that measuring HE4 in urine may aid diagnosis and the monitoring of response to therapy.

Keywords

ELISA; ovarian cancer diagnostics; urine; HE4

1. Introduction

There is a need for non-invasive assays to aid the early detection of ovarian carcinoma, and particularly its most aggressive serous form [1,2]. This has been a challenging task and it remains uncertain to what extent ovarian cancer can be detected when it is sufficiently small to be curable by existing therapies [3]. Assays measuring CA125 in serum have high sensitivity, but CA125 levels are frequently elevated in women with benign disease [4–6]. Furthermore, recent results from the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO) revealed that most postmenopausal women with an elevated serum CA125 detected during screening already had advanced cancer [7]. There is thus a need for additional markers that can complement CA125 both for early detection and to monitor the response to therapy and many such markers have been reported. One of them is mesothelin
[8,9], and a double determinant (“sandwich”) ELISA [10], primarily recognizing its variant 1 [11], complements CA125 [12,13] and also aids the diagnosis of mesothelioma [14].

The WFDC2 gene [15] is frequently amplified in ovarian carcinoma, whereas its expression in normal tissues, including ovary, is low [16–19] with the highest normal tissue expression in cells from glandular and respiratory epithelium [19,20]. WFDC2 encodes HE4, a member of a family of stable 4-disulfide core proteins, which include the Wp protein [21], a leukocyte protease inhibitor [22] and elafin [23], all of which are secreted molecules. An ELISA was constructed to measure HE4 and used to test sera from postmenopausal women with ovarian cancer [24]. It was shown to have a sensitivity equivalent to that of CA125 and was positive less frequently than CA125 in patients with nonmalignant disease [24]. These findings were confirmed and extended in a large, double-blind study which was performed at several academic centers on sera from women who had a pelvic mass, where tests with a commercially available kit showed that a combination of HE4 with CA125 identified a larger fraction of patients with ovarian cancer than CA125 alone [25,26]. Furthermore, a recent systematic evaluation of candidate blood markers for detection of ovarian cancer indicated that HE4, together with mesothelin, Muc16 and MMP7, is one of the most promising newly reported markers to identify women who have ovarian cancer [27].

Detection of biomarkers in urine can provide a less invasive and more convenient way of identifying ovarian cancer and may facilitate longitudinal studies of patients over a period of time to aid the early detection of patients with ovarian cancer as well to monitor responses to therapy. Several urine markers have been identified, including eosinophil-derived neurotoxin, a fragment of osteopontin [28], mesothelin [29], and Bcl-2 [30]. Since the HE4 protein has a molecular weight around 25 kD we hypothesized that it can be detected in urine and investigated whether testing of urine can distinguish patients with ovarian carcinoma from control groups with sufficient sensitivity to warrant further studies to learn whether such testing can complement or possibly replace the testing of serum.

2. Materials and Methods

2.1. Study population and sample collection

As listed in Table 1, urine samples were obtained, one sample/subject, from 36 age matched healthy women (used as normal controls), 20 patients with benign gynecological disease, 15 patients with early (stage I/II) and 64 patients with advanced (stage III/IV) ovarian cancer. 4 of the early stage cancers and 38 of the advanced ones were of the serous histology. All cases were surgically staged according to guidelines by the International Federation of Gynecologic Oncology (FIGO). Urines were tested from all subjects from whom samples were available in the tissue bank (headed by E.S.) at Dept. of Gynecology/Obstetrics, University of Washington. As shown in Table 1, most women were postmenopausal with an approximately equal distribution comparing controls and cases. Women with benign disease had a variety of pathological diagnoses including benign ovarian cystadenoma, uterine fibroids, endometriosis, physiological ovarian cysts and normal findings. We also tested a separate set of urine and serum samples which had been obtained concomitantly from 32 women with stage III/IV ovarian cancer.

Urine samples and clinical information were obtained from the University of Washington Gynecologic Oncology Tissue Bank as approved by the Human Subjects Division of the Institutional Review Board. All urine samples from patients were harvested during surgery for cases and women with benign disease. Normal controls donated a random urine collection. After collection, the samples were centrifuged and the non-cellular supernatant stored at −80°C. Specimen collection and processing protocols were identical for all women regardless of case or control status. All subjects were characterized at the time of specimen
collection with respect to age and menopausal status. Women were considered postmenopausal if they reported no menstruation for 6 months or were above 50 and did not report menstrual history. Stage and histology were recorded for all cancer cases. Samples were assayed for HE4 as well as for creatinine to relate the findings to the varying protein concentration in urine samples.

2.2. Enzyme-linked immunosorbent assay

HE4 levels were measured by a quantitative enzyme-linked immunosorbent assay (ELISA) provided as a kit by Fujirebio Diagnostics Inc (FDI), using the manufacturer’s instructions. Plates were read at 405 nm wavelength using a microplate reader (Fusion universal microplate analyzer, Fusion instrument company, Meriden, CT) within 15 min. Pilot tests showed that urine samples had to be diluted 1:40 prior to testing because of an otherwise high background; the dilution was made with tracer diluent (phosphate buffered saline with BSA, blocking agents, detergents, an inert blue dye and a non-azide antimicrobial preservative) as provided with the kit.

Urinary creatinine levels were measured using a kit (Cayman Chemical Company, Ann Arbor, MI) and following the manufacturer’s protocol. ELISA data from urine samples were normalized by calculating the ratio HE4 (pM)/creatinine (mg/dl) or mesothelin OD450/creatinine (mg/dl) where the mesothelin value was multiplied with 1,000.

2.3. Statistical analysis

All raw data were submitted to the group’s biostatistical expert (P.J.H.) for analysis. In order to evaluate the ability of the HE4 values to discriminate cases from controls we first used numerical summaries (mean and standard deviation) as well as graphical summaries such as box plots and scatter plots. Based on the skewed distribution of the measurements we chose to use a log transform (log10) rather than directly use the raw measurements. We analyzed the “level” of the marker defined as the log transformed quantification in urine, as well as the “ratio” of the marker defined as the log-transformation of the ratio of the marker value to the creatinine level. Therefore, we consider two total candidate markers: HE4 level and ratio. To explore the relationship between the marker “level” and marker “ratio” we used scatter plots with distinct plotting characters to denote each study group (controls, early cases, late cases). For each of the two candidate markers, ROC curves were computed to show the possible combinations of sensitivity and specificity associated with different choices for the cut-point used to classify marker values as a positive test. For each marker we computed two ROC curves: one ROC curve comparing early cases to controls, and a second ROC curve comparing late cases to controls. The area under the ROC curve (AUC) was calculated for each marker and bootstrap confidence intervals for the AUCs were calculated. We also used the ROC curves to estimate the sensitivity of the marker for detecting cases of each type by using a cut-off value that yields high specificity. In particular we computed sensitivity associated with a test-positive threshold where only 2/36 controls test positive, i.e. where specificity is 34/36 = 94.4%. Again we used bootstrap methods to compute confidence intervals for the sensitivity associated with the fixed 94.4% level of specificity.

We also explored whether markers could be combined to increase their discriminatory potential. Rather than directly try to combine both markers using the relatively small sample size we chose to derive a linear combination of HE4 level and ratio (HE4 combo marker). In order to compute the linear combination of markers that would predict disease status we used logistic regression using only the early cases and the controls. Specifically, we used the model: \[ \log \left( \frac{p}{1-p} \right) = b_0 + b_1 \cdot \log10(\text{level}) + b_2 \cdot \log10(\text{ratio}), \] where \( p \) = probability of \( Y=1 \) using \( Y=0 \) to denote a control subject and \( Y=1 \) to denote a case subject. The result from
logistic regression is a derived predictive marker given by $b_1 \times \log_{10}(\text{level}) + b_2 \times \log_{10}(\text{ratio})$. Rather than use the exact estimated values for $b_1$ and $b_2$ we used an approximate value by normalizing the values to be $w_1 = b_1 / (b_1 + b_2)$ and $w_2 = b_2 / (b_1 + b_2)$ such that the derived marker would be a simple weighted average of the component measures, and then we rounded these estimated weights to two significant digits. We then computed the derived marker for HE4 and evaluated its ability to separate cases from controls by computing ROC curves. Since we have used the data to derive the linear combination we used cross-validation in order to compute valid ROC curves [31]. Finally, we considered whether additional improvement could be obtained by using the two markers in a logistic regression for early case status.

### 3. Results

Figure 1 shows the HE4 in urine data as a scatter plot for controls (36 healthy women), 15 patients with early (stage I/II) and 64 patients with late (stage III/IV) ovarian carcinoma. All raw data are included with antigen levels (x-axis) and ratios between antigen levels and creatinine (y-axis). Both the levels and ratios were low in the control samples as compared to samples from patients with ovarian cancer, including patients with early disease.

Figure 2 presents box plots for the HE4 combo marker derived by combining both the level and ratio. The logistic regression analysis suggested that equally weighting the $\log_{10}$ (HE4 level) and the $\log_{10}$ (HE4 ratio) was optimal for predicting early case status versus control status. Specifically, the regression coefficient for $\log_{10}$ (HE4 level) was 4.53 (s.e.=2.37, $p=0.057$) and for $\log_{10}$ (HE4 ratio) was 12.50 (s.e.=5.06, $p=0.0135$) leading to weights $w_1=4.53/(12.50+4.53) = 0.27$ and $w_2=12.50/(12.50+4.53)=0.73$. Using this combined marker, the derived level of HE4 in the urine from patients with both early and late stage ovarian cancer was significantly higher than in the control group. Using logistic regression to predict late cases versus controls suggests a different weighted average for combining level and ratio. The estimated coefficient for $\log_{10}$ (HE4 level) as a predictor of late case status was $-1.06$ (s.e.=1.37, $p=0.439$) while the estimated coefficient of $\log_{10}$ (HE4 ratio) was 17.06 (s.e.=4.61, $p<0.001$). These results suggest that HE4 ratio is the major predictor of both early and late case status. Since we prioritized the prediction of early cases we chose to use weights of $w_1=0.73$ for $\log_{10}$ (HE4 ratio) and $w_2=0.27$ for $\log_{10}$ (HE4 level) when deriving a combined HE4 marker. It is noteworthy that the HE4 combo marker in urines from women with benign gynecological disease (G in Figure 2) is not significantly different from the combo marker for urine from the control (healthy women). The derived HE4 marker (average of $\log_{10}$ level and ratio) predicted both early ($p=0.01$) and late ($p<0.001$) case status.

Figure 3 presents ROC curves for the HE4 combo marker, illustrating that HE4 can be detected with high specificity and sensitivity in urine. The area under the ROC curve, or AUC, is an overall summary of discrimination. For the derived HE4 marker the AUC for early cases is 0.969 (95% confidence interval 0.910 – 1.0) while for late cases the AUC is 0.964 (95% confidence interval 0.924 – 0.992). These AUCs indicate excellent diagnostic performance of urine HE4.

Cut-off values were calculated, as described above, to classify the data as positive or negative. Table 2 summarizes the data as defined by the combination markers and calculated at 94.4% specificity (2 of 36 healthy women positive). As shown in the table, HE4 was positive for 86.6% for patients with early disease (stage I and stage II) and 89.0% of patients with late (III/IV) stage. 3 of 20 patients with benign gynecological disease (15%) tested positive. If, instead, we applied a lower cut-off so that the combination marker for none of the 36 healthy women and none of the 20 women with benign gynecological disease was...
positive (100% specificity), 11 of 15 (73.3%) women with stage I/II and 56 of 64 (87.5%) women with stage III/IV disease were positive.

The serous histological type is both its most aggressive and most common variant of ovarian cancer [2]. It is noteworthy, that, as shown in Table 2, 38 of 42 patients with serous cancer (90.5%) had urines that were positive for HE4, including all of 4 patients with serous stage I/II ovarian cancer.

We did not have available samples of both urine and serum from the patients tested in the studies described above to allow comparisons between measuring HE4 in both urine and serum. However, we had access to a group of concomitantly obtained urine and serum samples from 32 additional patients with stage III/IV ovarian cancer. We tested those samples, using the cut-off established for the HE4 kit when analyzing serum and the cut-off for the HE4 combo marker to yield 94.4% specificity (as described above) when analyzing urine. With those criteria, 24 of 32 sera were positive as compared to 26 of 32 urines, indicating that assaying urine and serum gives comparable results.

4. Discussion

Our objective was to investigate whether HE4, one of the most promising new biomarkers for ovarian cancer [19,24,25,27], can be detected in urine as a potentially useful diagnostic tool. HE4 levels were analyzed both as the absolute level of antigen and the ratio between antigen level and urinary creatinine as a correlate of urine concentration which varies dependent on many factors, including when the urine is harvested. Since both measurements appeared to be informative, a combined (HE4 combo) marker was constructed as described above by taking into account both the level and the ratio. At 94.4% of specificity (2 of 36 samples from age matched healthy women positive), the urine test was positive in 86.6% of women with early (stage I/II) and in 89% of women with stage III/IV ovarian cancer. It is noteworthy that urines from 38 of 42 women with serous ovarian cancer, its most aggressive form, were positive. In contrast, there was only a small (and not statistically significant) increase in the frequency of positive urines from women with benign gynecological disease, which was at a level similar to that reported for serum [24]. At a cut-off when none of 36 healthy women or 20 women with benign gynecological disease had urine positive for HE4, 11 of 15 (73.3%) stage I/II and 56 of 64 (87.5%) stage III/IV patients were positive.

This is the first report indicating that measuring HE4 in urine is as sensitive and specific as reported from studies with serum published by ourselves [24] and others [13,25,26]. Although we have provided confidence intervals for all primary summary statistics, our studies are retrospective and performed on a relatively small number of controls and ovarian cancer patients of whom only 15 had early stage disease including 4 (albeit all positive for HE4) of the serous type for which the diagnostic need is highest.

Serial studies on the same patient can aid the early diagnosis of ovarian cancer, as demonstrated for CA125 [32]. Like CA125, HE4 has longitudinal stability [24], suggesting that frequent serial studies can make possible an earlier detection of ovarian cancer as well as relapse after treatment. Measuring HE4 in urine provides a noninvasive approach to perform such studies with a sensitivity and specificity, which is similar to that by assaying serum.

We feel that a prospective study is warranted to validate and expand our findings with a much larger sample size and focusing on patients with early ovarian carcinoma, particularly of the serous type. An analogous prospective study confirmed and expanded our original observation [24] of increased levels of HE4 in serum from patients with ovarian cancer [25,26]. An important objective of an expanded study will be to investigate the extent to
which the HE4 marker relates to tumor size which may be a better correlate of an aggressive ovarian cancer than its stage [3].

**Abbreviations**

- **ROC**: receiver operating characteristics
- **ELISA**: enzyme-linked immunosorbent assay

**Acknowledgments**

The work was supported by a grant from Fujirebio Diagnostics, Inc and by NIH grants 1RO1CA1344867 and UL1 RR025014. The support of Dr Nancy Kiviat is gratefully acknowledged.

**References**


Figure 1.
Urinary HE4 levels and HE4/creatinine ratios in ovarian cancer patients. The scatter plot shows increased levels (x-axis) and ratios (y-axis) in the urines of patients with early (stage I and II) and late (stage III and IV) ovarian cancer as compared to controls.
Figure 2.
Urinary HE4 levels are significantly elevated in ovarian cancer patients. The box plot shows that (HE4 combo) algorithms combining the HE4 level and HE4/creatinine ratio were significantly increased in the urines of women with early (I/II) and late (III/IV) ovarian cancer as compared to healthy women (H) and women with benign gynecological disease (G).
Figure 3.
A receiver operating characteristics (ROC) curve of HE4 in urine. It shows the relation between specificity and sensitivity for an algorithm (HE4 combo) combining the HE4 level and the HE4/creatinine ratio for women with early (stage I/II) and late (III/IV) ovarian cancer.
# Table 1

Summary description of study subjects who donated specimens for our analysis

<table>
<thead>
<tr>
<th>Menopausal status</th>
<th>Stage</th>
<th>OvC cases *</th>
<th>Benign gynecologic disease **</th>
<th>Healthy women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Premenopausal</td>
<td></td>
<td>23</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Postmenopausal</td>
<td></td>
<td>56</td>
<td>10</td>
<td>24</td>
</tr>
</tbody>
</table>

* includes 42 serous ovarian cancer, 15 adenocarcinoma, 7 endometroid, 4 clear cell, 3 mucinous, 8 others

** includes 2 endometriosis, 2 serous cystadenoma, 2 mucinous cystadenoma, 14 others
Table 2
Percentage urine samples, one per donor, which are positive according to the indicated cut-offs for the respective combo markers; number of positives over number tested is shown within parentheses.

<table>
<thead>
<tr>
<th></th>
<th>HE4 combo*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy women</td>
<td>5.6 (2/36)</td>
</tr>
<tr>
<td>Benign gynecologic</td>
<td>15.0 (3/20)</td>
</tr>
<tr>
<td>disease</td>
<td></td>
</tr>
<tr>
<td>Stage III ovarian</td>
<td>86.6 (13/15)</td>
</tr>
<tr>
<td>cancer, all patients</td>
<td></td>
</tr>
<tr>
<td>serous</td>
<td>100.0 (4/4)</td>
</tr>
<tr>
<td>Stage III/IV ovarian</td>
<td>89.0 (57/64)</td>
</tr>
<tr>
<td>cancer, all patients</td>
<td></td>
</tr>
<tr>
<td>serous</td>
<td>89.0 (34/38)</td>
</tr>
</tbody>
</table>

* Cut-offs were selected as described in Materials and Methods so that 2/36 control samples from healthy women were positive (specificity 94.4%).