Abstract. Aim: A nested case-control discovery study was undertaken to test whether information within the serum peptidome can improve on the utility of CA125 for early ovarian cancer detection. Materials and Methods: High-throughput matrix-assisted laser desorption ionisation mass spectrometry (MALDI-MS) was used to profile 295 serum samples from women pre-dating their ovarian cancer diagnosis and from 585 matched control samples. Classification rules incorporating CA125 and MS peak intensities were tested for discriminating ability. Results: Two peaks were found which in combination with CA125 discriminated cases from controls up to 15 and 11 months before diagnosis, respectively, and earlier than using CA125 alone. One peak was identified as connective tissue-activating peptide III (CTAPIII), whilst the other was putatively identified as platelet factor 4 (PF4). ELISA data supported the down-regulation of PF4 in early cancer cases.

Conclusion: Serum peptide information with CA125 improves lead time for early detection of ovarian cancer. The candidate markers are platelet-derived chemokines, suggesting a link between platelet function and tumour development.

There are over 220,000 new cases of ovarian cancer worldwide each year (1). The disease has a poor prognosis mainly due to late diagnosis, with over 70% of patients exhibiting spread beyond the pelvis (2). Currently, ovarian cancer screening is not recommended due to lack of evidence of a mortality benefit, although large screening trials are underway to explore this issue in the low-risk population (3, 4). Strategies being tested combine serum cancer antigen 125 (CA125) assay with transvaginal sonography (TVS). CA125 is a coelomic epithelium-related glycoprotein of unknown function, secreted into the bloodstream by ovarian epithelial cells. However, CA125 is also produced by other mesothelium-derived tissues and may thus be elevated in benign gynaecological conditions and non-ovarian carcinomas, whilst not all early-stage tumours generate CA125 (5). Consequently, the diagnostic accuracy of CA125 using a cut-off is sub-optimal and additional biomarkers to improve accuracy for screening and early detection of ovarian cancer are required (6).

The performance characteristics required of an ovarian cancer biomarker depend upon the intended clinical use. Markers that can differentiate benign pelvic masses from malignancy in symptomatic women and guide surgical decisions need to improve upon existing tests which achieve sensitivities of 85-90% for detecting symptomatic ovarian cancer and specificities of 85-90% for benign disease (7). Indeed, one such multi-marker test OVA1 (which includes CA125) was recently approved by the FDA (8). When combined with clinical assessment by imaging and physical examination, it achieves a sensitivity of >90% and a negative predictive value of 90% in women with an ovarian tumour. A more challenging clinical use is screening for ovarian cancer. This requires high sensitivity together with specificity in excess of 98% so that at least one in ten women who undergo surgery as a result of screen positive results (positive predictive value (PPV) >10%) has ovarian cancer. In the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) this was achieved in the prevalence screen by using CA125 interpreted using the risk of ovarian cancer (ROC) algorithm followed by TVS (3). While it is
encouraging that 48% of the cases detected were early-stage cancer, it does raise the need for improving lead time even when high sensitivity and specificity are achieved.

Mass spectrometry (MS)-based proteomic profiling of polypeptides in serum (or plasma) holds promise for the identification of novel cancer biomarkers (9-11). However, serum proteome profiling is challenging due to the complexity and large dynamic range of abundance of its component proteins. Together with significant intra- and inter-individual heterogeneity, this has meant that few, if any, robust cancer biomarkers have been identified to date using proteomic methods. Indeed, most candidates reported have been abundant, non-specific acute-phase proteins (12, 13) that are likely to be secondary host responses to any diseased state rather than specific markers useful for accurate diagnosis (14). In addition, numerous studies have highlighted alterations in serum and plasma proteins that are attributable to sample handling that is largely driven by differential proteolysis (15-21) (22). Finally, concerns have been raised over assay reproducibility and the robustness of class-discriminating algorithms used for proteomic profiling biomarker discovery (14, 23, 24).

Matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) MS is a technique that can be used for high-throughput profiling of clinical samples, particularly when linked to automated handling. MALDI-TOF MS profiling has revealed the complexity of the low-molecular weight proteome (peptidome) of serum and plasma (25-27). Most peptides observed in MALDI-TOF profiles are derived from relatively few abundant proteins as the result of extensive proteolysis, and it has been reported that these cleavage patterns can discriminate cancer cases from healthy controls (28). It was hypothesized that these peptide patterns are generated by tumour-specific exopeptidases during coagulation, and as such, represent *ex vivo* surrogate markers of cancer (28).

In the present discovery study, we have used semi-automated peptide extraction linked to MALDI-TOF MS to profile the peptidomes of serum samples taken from the UKCTOCS trial. Samples were obtained from women at various timepoints prior to diagnosis of primary invasive ovarian cancer and from healthy controls which were matched based on clotting time, collection centre and age. Given that these samples pre-date symptomatic cancer, they provide a unique source of early markers in the absence of late-stage confounding markers. We tested the power of peptide peaks alone and in combination with serum CA125 for discriminating cases from controls at different timepoints prior to diagnosis building on previous work from a pilot study where we reported that combining a single MS peak with CA125 allowed significant discrimination between cases and controls up to 12 months in advance of diagnosis (29).

Table I. *Baseline characteristics of cases and controls with significance testing.*

<table>
<thead>
<tr>
<th>Baseline characteristics</th>
<th>Controls (n=585)</th>
<th>Cases (n=159)</th>
<th>Mann-Whitney test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years) at randomisation</td>
<td>59.8 (55.6-65.2)</td>
<td>61.7 (57.3-66.4)</td>
<td>0.01</td>
</tr>
<tr>
<td>Age at last period (years)</td>
<td>50.1 (46.8-52.7)</td>
<td>50.7 (46.2-52.6)</td>
<td>0.873</td>
</tr>
<tr>
<td>Duration of HRT use in those on HRT at randomisation (years)</td>
<td>7.8 (4.5-10.7)</td>
<td>8.9 (4.8-13.8)</td>
<td>0.263</td>
</tr>
<tr>
<td>Duration of OCP use (years) in those who had used it</td>
<td>5 (2-10)</td>
<td>5 (2-10)</td>
<td>0.639</td>
</tr>
<tr>
<td>Miscarriages (pregnancies &lt;6 months)</td>
<td>0 (0-1)</td>
<td>0 (0-1)</td>
<td>0.656</td>
</tr>
<tr>
<td>No. of children (pregnancies &gt;6 months)</td>
<td>2 (2-3)</td>
<td>2 (1-3)</td>
<td>0.167</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>162.6 (157.5-165.1)</td>
<td>162.6 (157.5-167.6)</td>
<td>0.809</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>67.1 (60.8-76.2)</td>
<td>68.7 (60.3-76.2)</td>
<td>0.601</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Number (%)</th>
<th>Fisher’s exact test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethnicity</td>
<td></td>
<td>0.59</td>
</tr>
<tr>
<td>White</td>
<td>571 (97.6%)</td>
<td>155 (97.5%)</td>
</tr>
<tr>
<td>Black</td>
<td>4 (0.7%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Asian</td>
<td>5 (0.9%)</td>
<td>2 (1.3%)</td>
</tr>
<tr>
<td>Other</td>
<td>2 (0.3%)</td>
<td>1 (0.6%)</td>
</tr>
<tr>
<td>Missing</td>
<td>3 (0.5%)</td>
<td>1 (0.6%)</td>
</tr>
<tr>
<td>Hysterectomy</td>
<td>101 (17.3%)</td>
<td>27 (17.0%)</td>
</tr>
<tr>
<td>Ever use of OCP</td>
<td>327 (55.9%)</td>
<td>69 (43.4%)</td>
</tr>
<tr>
<td>Use of HRT at recruitment</td>
<td>126 (21.5%)</td>
<td>41 (25.8%)</td>
</tr>
<tr>
<td>Maternal history of ovarian cancer</td>
<td>30 (5.1%)</td>
<td>5 (3.1%)</td>
</tr>
<tr>
<td>Maternal history of breast cancer</td>
<td>102 (17.4%)</td>
<td>31 (19.5%)</td>
</tr>
</tbody>
</table>

HRT, Hormone replacement therapy; OCP, oral contraceptive pill.
Materials and Methods

Serum samples. A nested case-control study was undertaken on serum samples collected from 159 women who developed primary invasive ovarian cancer and 585 healthy women recruited to the UKCTOCS (30) (http://www.instituteforwomenshealth.ucl.ac.uk/academic_research/gynaecologicalcancer/gcrc/ukctocs/design). This secondary study was approved by the Joint UCL/UCLH Research Ethics Committees on the Ethics of Human Research, Committee A (Ref. No. 05/Q0505/57) and written informed consent was obtained from all donors. No data allowing identification of patients was provided. There were 880 samples in total collected between April 2001 and January 2007 at 13 trial centres, of which 295 samples were from the 159 women who went on to develop ovarian cancer (referred to as cases). Of these there were single samples from 90 cases and up to 5 serial samples from 69 cases. According to a standard operating procedure used at all centres, blood samples were taken by venepuncture into gel tubes (8 ml gel separation serum tubes; Greiner Bio-One, Stonehouse, UK) and transported by courier at ambient temperature to the central laboratory. All samples received more than 56 h after venepuncture were discarded and repeat samples requested. The blood was centrifuged at 1500 g for 10 min, the serum separated and aliquoted into bar-coded straws which were frozen at –80°C and then transferred the next day to liquid nitrogen for long-term storage. For each case the time to diagnosis was known, and is the time interval (measured in months) between the date of collection and the date of diagnosis (defined as the date of surgery/biopsy). Matched controls were 585 samples from healthy women with two samples matched per case (only 5 cases had 1 matched control). These were taken as close as possible in time to the case sample on the same day and at the same collection centre, and hence were clotted for the same time and stored identically as matched cases. Clotting times ranged between 1 and 49 h (average 23 h ± 4 h). Samples were also matched by age at sample draw as a secondary criterion if time and centre matching allowed. Baseline characteristics of cases and controls and statistical assessment of differences are presented in Table 1.

Sample preparation, MALDI-TOF MS profiling and data processing. Blinded serum samples were subjected to pre-fractionation using robotic C18 ZipTip® (Billerica, MA, USA) reversed-phase extraction prior to automated MALDI-TOF MS data acquisition as described previously (31)-(32). Briefly, polypeptides were enriched from 5 μl of serum using a semi-automated protocol based on reversed phase pre-packed tips (C18 ZipTips®). A CyBio®-Disk robot (CyBio AG, Jena, Germany) equipped with a 96-piston head for 25 μl tips was adapted and used for this purpose. After C18 ZipTip® purification, enriched polypeptides were eluted from the ZipTips® and 2 μl from the eluate were mixed with α-cyano-4-hydroxycinnamic acid (CHCA) matrix and spotted onto a 600 μm-AnchorChip target plate (Bruker Daltonics, Bremen, Germany) for analysis on an Ultraflex II MALDI-TOF/TOF mass spectrometer (Bruker Daltonics) using FlexAnalysis v3.0 software (Bruker Daltonics) for further data handling. Up to 12 replicate mass spectra were generated for each sample giving 11,048 mass spectra in total. These were mass calibrated against peptide and protein standards run at the same time and converted to two column ASCII files of m/z values and corresponding intensities for further processing and analysis. Data processing steps (data reduction, smoothing, baseline subtraction, normalization, peak defining and peak alignment) were applied to the data, as previously described (29). Documentation and Matlab system code for these steps are available at http://www.clrc.rhul.ac.uk/projects/proteomic3.htm. Following processing, a set of peak groups were defined, ordered and numbered according to their frequency of appearance in the dataset. The 67 most common peaks appearing in >33% of all spectra (cases and controls) were used for subsequent data analysis (Table II). Finally, the intensities of the 67 peaks were averaged across replicates for each sample.

Classification and statistical testing. For classification, data from cases with only one matched control, and samples with a clotting time greater than 24 h were excluded, resulting in 179 samples (from 104 cases) and 358 matched control samples organized into 179 triplets. A total of 59 cases had one measurement, 26 had 2 measurements, 11 had 3 measurements, 5 had 4 measurements, and 3 had 5 measurements. Each triplet was assigned to a time to diagnosis group (in months) using a 6-month time window. Log-linear models using peak intensity were then tested for triplet classification, i.e. identification of the case sample among controls and cases) were used for subsequent data analysis (Table II).

### Table II. List of the 67 most frequently occurring MALDI-MS peaks in the dataset showing their peak number assignment, processed m/z values and frequency of occurrence.

<table>
<thead>
<tr>
<th>Peak number</th>
<th>m/z Value (Da)</th>
<th>Frequency (out of 11048)</th>
<th>Peak number</th>
<th>m/z Value (Da)</th>
<th>Frequency (out of 11048)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4213.4</td>
<td>11043</td>
<td>35</td>
<td>1704.5</td>
<td>7292</td>
</tr>
<tr>
<td>2</td>
<td>7772.1</td>
<td>11040</td>
<td>36</td>
<td>992.2</td>
<td>7109</td>
</tr>
<tr>
<td>3</td>
<td>9297.8</td>
<td>11040</td>
<td>37</td>
<td>2084.5</td>
<td>7048</td>
</tr>
<tr>
<td>4</td>
<td>5341.0</td>
<td>11036</td>
<td>38</td>
<td>9724.2</td>
<td>6658</td>
</tr>
<tr>
<td>5</td>
<td>5909.2</td>
<td>11029</td>
<td>39</td>
<td>5583.3</td>
<td>6570</td>
</tr>
<tr>
<td>6</td>
<td>6636.0</td>
<td>10994</td>
<td>40</td>
<td>4285.7</td>
<td>6517</td>
</tr>
<tr>
<td>7</td>
<td>4647.8</td>
<td>10904</td>
<td>41</td>
<td>2993.5</td>
<td>6460</td>
</tr>
<tr>
<td>8</td>
<td>4057.2</td>
<td>10878</td>
<td>42</td>
<td>2606.0</td>
<td>6422</td>
</tr>
<tr>
<td>9</td>
<td>3243.9</td>
<td>10739</td>
<td>43</td>
<td>2274.3</td>
<td>6358</td>
</tr>
<tr>
<td>10</td>
<td>3959.0</td>
<td>10705</td>
<td>44</td>
<td>740.9</td>
<td>6349</td>
</tr>
<tr>
<td>11</td>
<td>4967.9</td>
<td>10682</td>
<td>45</td>
<td>2512.8</td>
<td>6345</td>
</tr>
<tr>
<td>12</td>
<td>3772.6</td>
<td>10609</td>
<td>46</td>
<td>4759.2</td>
<td>6193</td>
</tr>
<tr>
<td>13</td>
<td>2757.3</td>
<td>10596</td>
<td>47</td>
<td>3527.0</td>
<td>6005</td>
</tr>
<tr>
<td>14</td>
<td>8610.9</td>
<td>10537</td>
<td>48</td>
<td>1742.6</td>
<td>5873</td>
</tr>
<tr>
<td>15</td>
<td>8937.3</td>
<td>10437</td>
<td>49</td>
<td>3161.7</td>
<td>5591</td>
</tr>
<tr>
<td>16</td>
<td>4478.8</td>
<td>10393</td>
<td>50</td>
<td>3886.1</td>
<td>5586</td>
</tr>
<tr>
<td>17</td>
<td>2662.9</td>
<td>10367</td>
<td>51</td>
<td>6385.4</td>
<td>5361</td>
</tr>
<tr>
<td>18</td>
<td>1547.6</td>
<td>10222</td>
<td>52</td>
<td>1262.6</td>
<td>5188</td>
</tr>
<tr>
<td>19</td>
<td>5007.8</td>
<td>10174</td>
<td>53</td>
<td>853.1</td>
<td>5001</td>
</tr>
<tr>
<td>20</td>
<td>2381.1</td>
<td>10088</td>
<td>54</td>
<td>1790.0</td>
<td>4910</td>
</tr>
<tr>
<td>21</td>
<td>3510.1</td>
<td>9870</td>
<td>55</td>
<td>6810.0</td>
<td>4909</td>
</tr>
<tr>
<td>22</td>
<td>6437.3</td>
<td>9819</td>
<td>56</td>
<td>7476.7</td>
<td>4859</td>
</tr>
<tr>
<td>23</td>
<td>2364.6</td>
<td>9749</td>
<td>57</td>
<td>905.9</td>
<td>4733</td>
</tr>
<tr>
<td>24</td>
<td>2025.7</td>
<td>9274</td>
<td>58</td>
<td>1041.2</td>
<td>4627</td>
</tr>
<tr>
<td>25</td>
<td>8135.7</td>
<td>9195</td>
<td>59</td>
<td>2191.2</td>
<td>4583</td>
</tr>
<tr>
<td>26</td>
<td>2116.6</td>
<td>9140</td>
<td>60</td>
<td>1618.2</td>
<td>4304</td>
</tr>
<tr>
<td>27</td>
<td>1946.9</td>
<td>9042</td>
<td>61</td>
<td>870.1</td>
<td>4298</td>
</tr>
<tr>
<td>28</td>
<td>1467.7</td>
<td>9024</td>
<td>62</td>
<td>2871.2</td>
<td>4294</td>
</tr>
<tr>
<td>29</td>
<td>2935.4</td>
<td>8984</td>
<td>63</td>
<td>3266.5</td>
<td>4191</td>
</tr>
<tr>
<td>30</td>
<td>1450.7</td>
<td>8545</td>
<td>64</td>
<td>6231.3</td>
<td>3841</td>
</tr>
<tr>
<td>31</td>
<td>1016.7</td>
<td>8432</td>
<td>65</td>
<td>1278.4</td>
<td>3724</td>
</tr>
<tr>
<td>32</td>
<td>2212.2</td>
<td>8034</td>
<td>66</td>
<td>3616.4</td>
<td>3683</td>
</tr>
<tr>
<td>33</td>
<td>1898.9</td>
<td>7797</td>
<td>67</td>
<td>1115.1</td>
<td>3650</td>
</tr>
<tr>
<td>34</td>
<td>3194.8</td>
<td>7599</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CA125 level and the intensity of peak p, respectively, and \( w \) and \( v \) are various weights \((w=\{0, 0.5, 1, 2\}; v=\{-1, 1\})\), the latter taking into account the direction of regulation of the peak in controls versus cases, logarithms were taken to remove the arbitrary units of measurements. Error rates were then calculated for various classification rules taking the peak giving the least number of errors. If two rules led to the same error rate, the rule involving the most frequently occurring peak across the dataset had priority when calculating \( p \)-values.

In order to check the robustness of the triplet classification, three types of statistical tests were used which reject the following null hypotheses about the classification. (i) Assignment of the label ‘case’ within each triplet of each time group is random. We apply this hypothesis to compare the performance of the classification rules on the actual data set with a randomly permuted data set. (ii) Assignment of labels within triplets is independent of CA125 levels – we wanted to check here if using CA125 alone is good enough to separate cases from controls. (iii) Peak intensities do not contain information useful to improve the predictive ability of CA125. Rejection of this third hypothesis means that separation between cases and controls is significantly improved after adding peak intensities to the information given by CA125. Based on the Monte-Carlo method, we developed a procedure to calculate \( p \)-values (see (29)) for the first null hypothesis. Here, the test statistic is calculated a large number of times \((N)\) for the data set with a randomly assigned label ‘case’ within each triplet, and counting the number of times \((Q)\) this statistic is equal to or less than the actual statistic computed from the observed data set. The \( p \)-value is then estimated as \( Q/N \). For the second null hypothesis, the error rate was computed using CA125 only \((i.e., v=0)\). For the third null hypothesis, the same set of classification rules is used as for the first null hypothesis, but instead of randomly assigning the ‘case’ label, we randomly permute the three peak intensities within each triplet \((6 \text{ possible permutations})\), leaving the labels and CA125 levels as they are. All three procedures produce valid \( p \)-values that do not need adjustment in the sense that for each threshold \( \delta \) the probability that the computed \( p \)-value does not exceed \( \delta \) will be at most \( \delta \) (under the null hypothesis).

A detailed account of the corresponding \( p \)-value calculations using the Monte-Carlo method, their meanings and validity is described at the end of the results section (see also (29)).

**Peak identification.** MALDI-TOF MS peaks were identified using 96 pooled UKCTOCS sample eluates from C18 ZipTip® extraction as a source of material and a combination of MALDI-MS/MS, ESI MS/MS and off-line LC fractionation and ESI MS/MS with or without tryptic digestion. The methodology (see (27)) was designed to ensure that identifications could be directly linked to the relevant peak by confirming the presence of that peak in fractions at each stage of the procedure by MALDI-TOF MS. For peak identification using MALDI-MS/MS, the pooled eluates were mixed in a ratio of 1:1 \((v:v)\) with either 5 mg/ml CHCA or 50 mg/ml 2,5-dihydroxybenzoic acid (DHB) matrix prepared in 50% acetoniitile (ACN)/0.1% trifluoroacetic acid (TFA) and spotted onto a MALDI target plate for analysis on a Q-TOF Premier (Waters, Manchester, UK) equipped with a 337 nm-UV-MALDI source. For peak identification using ESI MS/MS, the pooled eluates were first dried, then resuspended in 50% ACN/0.1% formic acid and sprayed using a Nanomate HD (Adwon Biosciences, Ithaca, NY, USA) as nanoESI ion source in front of the Q-TOF Premier. A positive voltage of 1.7 kV was applied to the chip and the flow rate was kept constant at ~80 nL/min. At this flow rate, a sample volume of 10 μL provided stable static electrospray for at least 2 h.

For peak identification using off-line LC fractionation and ESI MS/MS, 100 μL of the pooled eluate was loaded onto a HiChrom ACE C18 column \((2.1 \text{mm ID, } 150 \text{ mm length})\) (HiChrom Ltd., Reading, UK), pre-equilibrated with 5% ACN/0.1% formic acid (FA) and using an Agilent 1100 LC system (Agilent Technologies inc., Santa Clara, CA, USA). Elution was achieved using a binary solvent gradient of 5-70% ACN in 0.1% FA at a flow rate of 100 μL/min over 90 min, and finally up to 95% ACN/0.1% FA in 10 min. UV detection was set at 214 nm with fractions collected every minute. Fractions were dried and resuspended in 40 μL of 50% ACN/0.1% FA, then 15 μL was taken for chip-based nanoESI analysis using the Nanomate and Q-TOF Premier. 2 μL were used for MALDI-TOF analysis (Ultraflex), and the remaining was dried before resuspension in 15 μL of trypsin solution \((100 \text{ ng trypsin in } 100 \text{ mM ammonium bicarbonate, } pH 8.5)\) and incubation overnight at 37°C. Finally, digested fractions were subjected to C18 ZipTips® clean-up as detailed above for MALDI-TOF MS profiling, except that the elution was performed using 10 μL of 50% ACN/0.1% FA. The purified digested fractions were then loaded on the Nanomate system and nanoESI MS/MS data were collected using data dependent acquisition.

Identification of peptides was performed by searching against the human protein sequence library in NCBIInr (v20081128, 216937 sequences) using the Mascot 2.2 search engine (Matrix Science, London, UK). Searches were performed choosing “None” for enzyme (“Trypsin” for the digested fractions) and with a mass tolerance of 0.1 Da for parent ions and 0.2 Da for fragment ions. “Ammonia-loss (N-term C)”, “Deamidated (NG)”, “Dehydrated (N-term C)”, “Oxidation (M)”, “Phospho (ST)”, and “Phospho (Y)” were set as variable modifications. GPMAW software v7.10 (Lighthouse data, Odense, Denmark) was used to match accurate mass data obtained from undigested samples with theoretical peptide mass spectra of protein hits obtained from the MASCOT analysis of the MS/MS data of the digested fractions. Isotope pattern software v3.0 (Bruker Daltonics) was used for comparison of experimental isotopomer distributions with the theoretical distribution from putative peptide identifications and the UniProt and NCBI protein databases were used to extract additional information for substantiating or rejecting putative identities, e.g. information with regard to alternative splicing, pro-sequence cleavage, disulfide bridges and post-translational modifications.

**CA125 and PF4 immunoassays.** CA125 analysis was performed using an electro-chemiluminescence immunoassay on a Roche Elecsys 2010 analyser (Roche Diagnostics, Burgess Hill, UK). The assay uses monoclonal antibodies OC125 as the detection antibody and M11 as the capture antibody (Fujirebio Diagnostics; Oxford Biosystems, Oxford, UK). PF4 assays were performed on 173 of the case samples and 173 matched controls using an Asserachrom PF4 ELISA kit (Diagnostica Stago UK Ltd). Serum samples were diluted 1:2,100 v/v with dilution buffer and assayed according to the manufacturer’s instructions.

**Results**

**Peak discovery.** MALDI-MS profiling of a set of blinded ovarian cancer pre-diagnosis sera and matched controls from UKCTOCS was conducted to identify possible candidate markers for early detection. Raw MS spectral data were processed (Figure 1A) and classification rules applied to peak
intensity information and CA125 values and these tested for
significance. Given the limited number of samples, 6-month
time slots starting in different months with cases and two
matched controls grouped into triplets were considered. For
each time slot, hypotheses of random label distribution were
tested, calculating p-values and looking for peaks carrying
information for discriminating pre-diagnosis cancer cases
from matched controls. Using single peak information alone,
several of the 67 peaks analysed (Table II) were able to
discriminate cases from controls at a confidence level of
p<0.05, and mostly at the early time slots (Table III).
However, none provided significant discrimination after
adjustment for multiple testing.

Peak information was next combined with CA125 values
to examine if any peaks could improve on CA125 in terms
of early detection. The expected probability of error in triple
classification is 2/3, so misclassifying 16 out of 30 triplets
(as for the 13 month time group) was not significant
(p=0.09), but still better than random (Table IV). The results
for classification using CA125 alone were significant up to
9 months before diagnosis, an important finding. When peak
information was added, the classification was improved,
with the lead time of detection significant (p<0.05) up to 15
months prior to diagnosis, with the only exception being the
12 months time group with a p-value of 0.059. Peaks with
processed m/z values of 7772 and 9297 (referred to as peaks
2 and 3 based on their frequency of occurrence) were used
most often for discrimination of cases and controls at these
earlier time slots. The spectral profiles of these peaks are
shown in Figure 1B. None of these p-values require
adjustment as they satisfy the property of validity described
in the Materials and Methods. However, they do not
demonstrate the significance of individual peaks; they only
show that the peaks are significant en masse. A further p-
value was thus calculated. Whereas the overall ‘main’ p-
value tests the hypothesis that CA125 and peak intensity do
not help to discriminate cases from controls, the ‘conditional’ p-values presented in Table IV test the hypothesis that given CA125, the peaks do not carry additional information useful for discrimination. For this analysis, the conditional p-values are more important, since it is known that CA125 is a useful biomarker and therefore the interest is in whether addition of other data leads to significant improvement. The conditional p-values show that the contribution of adding MS peaks becomes essential only from 10 to 15 months.

The significance of peaks 2 and 3 were next checked
directly by adjusting the main and conditional p-values
using the sets of rules with these peak numbers built in.
Table V presents the results for prediction by CA125 and
peak 2 (columns 5-8), and by CA125 and peak 3 (columns
9-12). Error rate and p-value for prediction using CA125
alone are included in the table for comparison. The ‘rule’
columns show the best rules selected. As above, the
conditional p-values show that improvement on CA125
cannot be achieved by adding information from either peak
2 or 3 up to 9 months before diagnosis, but that these peaks
do improve its predictive ability beyond 9 months. Indeed,
the results are significant for up to 15 months using peak
2 and 11 months using peak 3. The main and conditional
p-values given in Table V have been adjusted for multiple
hypothesis testing (see (33) and Lemma 2 at the end of the
results section). Figure 2 illustrates the performance of the
best classification rules log C − 2 log I(2) and log C − log
I(3), respectively, in comparison with the performance of
log C. The horizontal axis shows time to diagnosis, the
vertical axis triplet groups in this time interval. The figures
demonstrate that most triplets with the measurement date
close to diagnosis date are predicted correctly even by the
log C rule and that most samples where addition of a peak
to CA125 allows correct classification (marked as up-
directed triangles) are in the interval of 13-16 months prior
to diagnosis. Figure 3 shows the median dynamics of these
rules for case measurements. For each time moment, the
latest available case measurement for each triplet group is
taken into account. These measurements are averaged by
median through all triplet groups. The figures illustrate that
the values from the rules combining CA125 with peak
intensity are elevated earlier than when using CA125 alone
and that this is followed by the exponential growth of
CA125 closer to diagnosis. Notably, combining both peaks
with CA125 did not improve the accuracy compared to the
single peak models.

Peak identification. Identification of peak 3 (m/z 9297) was
achieved using a thoroughly executed fractionation and MS-
based analysis of UKTOCS samples, ensuring that the peak
of interest was retained after each preparation step (27). Briefly,
MALDI-MS spectra were acquired before fractionation and for
each fraction. These were compared with ESI MS spectra for
all fractions, making sure that the major peaks also appear as
major peaks within the respective ESI MS spectra. The
fractions in which peak 3 was obtained were also digested and
further analysed by MS/MS. From this analysis, three peptides
were obtained that identified connective tissue-activating
peptide III (CTAPIII) as the major component of the relevant
fractions. CTAPIII is a bioactive cleavage product of platelet
basic protein/C-X-C motif chemokine 7. As also both MALDI
and ESI MS of the undigested fractions revealed that the
peak(s) close in mass to CTAPIII account for the main ion
signal intensities in these spectra, it can be concluded that
within these fractions, the MALDI-MS peak 3 is in fact
CTAPIII. Figure 1 C, D, E, F show example spectra acquired
throughout the analysis. In addition, using GPMAW software
for matching sequences to m/z values obtained from the ESI
MS spectra, we found only CTAPIII (full sequence) to fit the
obtained MS peak. For further confirmation, we also generated the theoretical isotopic distribution of CTAPIII which matched closely to that found experimentally. Despite intensive effort, peak 2 (m/z 7772) eluded identification using these methods. Instead, we have relied on literature searches and the fact that the peak is relatively isolated in spectra and frequently occurring (Table II). The identity of peak 2 as platelet factor 4 (PF4/CXCL4; average molecular weight 7769.18 Da) was thus inferred from two studies where it was identified by offline LC fractionation and ESI MS/MS following SELDI-TOF MS serum profiling (34) and by immuno-capture after MALDI-TOF MS serum profiling (35). Although inferred, it is highly likely that peak 2 in our samples represents the chemokine PF4.

Figure 1. A: Processed full-mass range MS spectra for example case (red) and healthy control (green) samples. The spectra are averaged from 10 replicate acquisitions per sample. Peak 2 (m/z 7772) and peak 3 (m/z 9297) are indicated by arrows. B: Processed MS data for peak 2 (m/z 7772) and peak 3 (m/z 9297) plotted for all samples. Red is controls, blue is cases. C: MALDI MS spectrum of one of the fractions in which peak 3, identified as connective tissue-activating peptide III (CTAPIII), was eluted. The arrow indicates peak 3. D: Overlay of the spectrum in Figure S1A with randomly selected raw MALDI MS profiles from UKCTOCS serum analysis. E: ESI MS spectrum of the same sample used for the analysis that generated the spectrum shown in Figure 1C. F: m/z 1325-1333 region of the spectrum shown in Figure 1E. Single but not double oxidation is clearly visible, in agreement with CTAPIII possessing a single methionine residue.
Figure 1. continued
Figure 1. continued
Verification. In an attempt to verify the PF4 data, an ELISA was used to measure its levels in 173 of the case samples and 173 matched controls, and associations with case control status, time to diagnosis, clotting time, tumour stage, age, hormone replacement therapy (HRT) use and CA125 level were examined. There was no significant difference when all cases and controls were compared with median values of 6,958 IU/ml (range=1.642-15.096) for cases and 6.847 IU/ml (range=1.610-15.223) for controls (p=0.69). However, PF4 levels were significantly lower in cases in the 6-12 and 12-24 months time to diagnosis groups versus those in the 0-6 months time group, with p-values of 0.037 and 0.012, respectively (Figure 4A). There was also a trend for lower PF4 levels in cases versus controls in the distant time groups, but these were not significant at the 95% confidence level. There was no difference in PF4 level with HRT use, age or clotting time, whilst samples from cases with stage IV tumours at diagnosis had lower levels of PF4 versus those with other stages (e.g. p=0.027 versus stage I) and controls (e.g.
Table III. Statistical analysis of peaks for discrimination of cases and controls. P-values are presented for each of 67 peaks distributed by time-to-diagnosis time slots (t+6 months prior to diagnosis). P-values of <0.05 are indicated in bold italics. For adjusted P-values a threshold of 0.05/67=0.00075 should be considered as significant at the 95% confidence level.

<table>
<thead>
<tr>
<th>t</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nos. cases</td>
<td>68</td>
<td>56</td>
<td>47</td>
<td>36</td>
<td>27</td>
<td>23</td>
<td>20</td>
<td>17</td>
<td>17</td>
<td>20</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>30</td>
<td>25</td>
<td>20</td>
</tr>
</tbody>
</table>

Peak 1
- No. cases: 68
- P-values:
  - 0.507
  - 0.427
  - 0.567
  - 0.515
  - 0.330
  - 0.722
  - 0.080
  - 0.192
  - 0.467
  - 0.462
  - 0.471
  - 0.330
  - 0.722
  - 0.080

0.05/67=0.00075 should be considered as significant at the 95% confidence level.

Statistical analysis of peaks for discrimination of cases and controls. P-values are presented for each of 67 peaks distributed by time-to-diagnosis time slots (t+6 months prior to diagnosis). P-values of <0.05 are indicated in bold italics. For adjusted P-values a threshold of 0.05/67=0.00075 should be considered as significant at the 95% confidence level.
Finally, there was a clear rise in CA125 in cases in the lead up to diagnosis (Figure 4B), whereas PF4 showed no consistent change (Figure 4C). There was no correlation between PF4 and CA125 levels for either the case or control groups (Figure 4D), although examination of only samples with low CA125 (<30 U/ml) did reveal lower PF4 levels in cases versus controls that approached significance ($p=0.065$).

**Statistical analysis of all peaks**

In order to check the robustness of our triplet classification using MS peaks, we designed three types of statistical tests which reject the following null hypotheses about classification of $S_a$:

1. The null hypothesis that the assignment of labels within each triplet in $S_a$ is independent of the information contained in CA125 and all MS peaks.

2. The null hypothesis that the assignment of labels within triplets is independent of the CA125 levels.

3. The null hypothesis that the peak intensities do not contain any information useful for improving the predictive ability of CA125.

The detailed explanation of corresponding p-value calculation and their meaning is represented below:

- **The main p-value** (or just 'p-value') is based on the null hypothesis that the assignment of labels within each triplet in $S_a$ is independent of the information contained in CA125 and all MS peaks.

We calculated these and all other p-values using the Monte-Carlo method. Set $E_0 := \min_{p, w, v} \text{Err}(S_a; p, w, v)$ and $p_0 := \min\{p : 3\text{e}^{-3\text{e}}\text{Err}(S_a; p, w, v) = E_0\}$, that is, $E_0$ is the minimum number of errors occurring in prediction by CA125 and one of the peak intensities and $p_0$ is the peak with the highest commonality that can provide this minimum error rate.

The statistic equal to the pair $(E', p')$ calculated the same way as $(E_0, p_0)$ is collected for a large number $N$ (we used $N=10^8$) of times on the same data set with randomly permuted labels within each triplet.

We then calculate the number $Q$ of times the statistic is as good as or better than the statistic $(E_0, p_0)$ computed from the true labels, where 'as good as or better than' (denoted by $\leq$) is understood in the sense of the lexicographical order: $(E', p') \leq (E_0, p_0)$ means that either $E' \leq E_0$ or $E' = E_0$ and $p' \leq p_0$. The p-value is then estimated as $(Q+1)/(N+1)$.

**Algorithm 1: Main p-value calculation**

**Input:** $t$, time to diagnosis.

**Input:** $N = 10^8$, number of trials.

1. $E_0 := \min_{p, w, v} \text{Err}(S_a; p, w, v)$
2. $p_0 := \min\{p : 3\text{e}^{-3\text{e}}\text{Err}(S_a; p, w, v) = E_0\}$
3. $Q := 0$
4. for $j := 1, \ldots, N$
5. Assign the 'case' label to a randomly chosen sample in each triplet in $S_a$.
6. Assign the 'control' label to all other samples.
7. Recalculate $E' := \min_{p, w, v} \text{Err}(S_a; p, w, v)$
8. Recalculate $p' := \min\{p : 3\text{e}^{-3\text{e}}\text{Err}(S_a; p, w, v) = E'\}$
9. if $(E', p') \leq (E_0, p_0)$ then $Q := Q + 1$
10. end if
11. end for
12. Output: $(Q+1)/(N+1)$ as the p-value.

- The previous p-value demonstrates significance of predictions provided by CA125 levels in combination with peak intensities. To compare the discriminating ability of the combination of CA125 and peaks with the well-known benchmark, CA125 on its own, we need to calculate CA125 p-values. This test checks the null hypothesis that the labels at $S_a$ are independent of the CA125 levels.

The p-values were also calculated using the Monte-Carlo method. However, similar p-values can also be calculated theoretically, as described in Appendix A.

Suppose $E_0 = \text{Err}(S_a; 1, 1, 0)$; that is, $E_0$ is the number of errors occurring in prediction by CA125 only. The statistic equal to the number of errors is collected for $N = 10^8$ times on the same data set with randomly reassigned labels within each triplet, and then the algorithm counts the number $Q$ of times the statistic is as good as or better than the statistic $E_0$ computed from the true labels. The p-value is then estimated as $(Q+1)/(N+1)$.

**Algorithm 2: CA125 p-value calculation**

**Input:** $t$, time to diagnosis.

**Input:** $N = 10^8$, number of trials.

1. $E_0 := \text{Err}(S_a; 1, 1, 0)$
2. $Q := 0$
3. for $j := 1, \ldots, N$
4. Assign the 'case' label to a randomly chosen sample in each triplet in $S_a$.
5. Recalculate $E' := \text{Err}(S_a; 1, 1, 0)$
6. if $E' \leq E_0$ then $Q := Q + 1$
7. end if
8. end for
9. Output: $(Q+1)/(N+1)$ as the p-value.
Figure 4. A: Scatter dot plots for serum levels of platelet factor 4 (PF4) measured by ELISA for cases and matched controls in groups with different times to diagnosis. The horizontal bars indicate mean values. Significant changes between case groups are indicated. B: Continuous time to diagnosis data plotted for CA125. LOWESS curve fitting was applied to the CA125 data (solid line) C: Continuous time to diagnosis data plotted for PF4. D: PF4 and CA125 data were plotted against one another (for cases only). Linear regression curve fitting was applied to the PF4/CA125 plot (dashed line).
### Table IV. Initial statistical analysis of CA125 and MS peak information.

<table>
<thead>
<tr>
<th>t</th>
<th>No. of cases</th>
<th>No. of errors</th>
<th>CA125 p-value</th>
<th>CA125 and all peaks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>68</td>
<td>2</td>
<td>0.0001</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>56</td>
<td>4</td>
<td>0.0001</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>47</td>
<td>6</td>
<td>0.0001</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>36</td>
<td>8</td>
<td>0.0001</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>27</td>
<td>7</td>
<td>0.0001</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>23</td>
<td>7</td>
<td>0.0008</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>6</td>
<td>0.0001</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>17</td>
<td>6</td>
<td>0.0071</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>5</td>
<td>0.0021</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>20</td>
<td>7</td>
<td>0.0042</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>28</td>
<td>14</td>
<td>0.0503</td>
<td>6</td>
</tr>
<tr>
<td>11</td>
<td>28</td>
<td>15</td>
<td>0.1028</td>
<td>8</td>
</tr>
<tr>
<td>12</td>
<td>28</td>
<td>17</td>
<td>0.3164</td>
<td>10</td>
</tr>
<tr>
<td>13</td>
<td>30</td>
<td>16</td>
<td>0.0895</td>
<td>12</td>
</tr>
<tr>
<td>14</td>
<td>25</td>
<td>16</td>
<td>0.4661</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>20</td>
<td>13</td>
<td>0.5211</td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td>10</td>
<td>6</td>
<td>0.4406</td>
<td>2</td>
</tr>
</tbody>
</table>

- **Columns in order are:** time \( t \) to diagnosis in months; the number of cases (triplets) with measurement taken between \( t \) and \( t + 6 \) months prior to diagnosis; the number of errors when classifying triplets with CA125 alone and associated \( p \)-value; the minimal number of errors when classifying with \( w \log C + v \log I(p) \); most often used peak (see Table II); the ratio of weights \( v/w \) used in the rule; the main \( p \)-value for overall significance of this result; the conditional \( p \)-value for significance of non-CA125 component. Bold areas show significance at \( p \)-value <0.05 for CA125 and main \( p \)-values beyond 9 months.

### Table V. Experimental results for triplet classification with a fixed peak (peak 2 or 3) and CA125. Results for prediction by CA125 alone (columns 3-4), CA125 and peak 2 (columns 5-8) or CA125 and peak 3 (columns 9-12) are shown for each time to diagnosis group.

<table>
<thead>
<tr>
<th>t</th>
<th>No. of cases</th>
<th>Error b</th>
<th>CA125 p-value</th>
<th>CA125 and peak 2</th>
<th>CA125 and peak 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>68</td>
<td>2</td>
<td>0.0001</td>
<td>2-CA-p02 2</td>
<td>2CA-p03 2</td>
</tr>
<tr>
<td>1</td>
<td>56</td>
<td>4</td>
<td>0.0001</td>
<td>2-CA-p02 4</td>
<td>2CA-p03 4</td>
</tr>
<tr>
<td>2</td>
<td>47</td>
<td>6</td>
<td>0.0001</td>
<td>2-CA-p02 5</td>
<td>2CA-p03 5</td>
</tr>
<tr>
<td>3</td>
<td>36</td>
<td>8</td>
<td>0.0001</td>
<td>2CA-p02 7</td>
<td>2CA-p03 7</td>
</tr>
<tr>
<td>4</td>
<td>27</td>
<td>7</td>
<td>0.0001</td>
<td>2CA-p02 6</td>
<td>2CA-p03 6</td>
</tr>
<tr>
<td>5</td>
<td>23</td>
<td>7</td>
<td>0.0008</td>
<td>CA-p02 6</td>
<td>CA-p03 6</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>6</td>
<td>0.0001</td>
<td>CA-p02 5</td>
<td>CA-p03 5</td>
</tr>
<tr>
<td>7</td>
<td>17</td>
<td>6</td>
<td>0.0071</td>
<td>CA-p02 4</td>
<td>CA-p03 5</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>5</td>
<td>0.0021</td>
<td>CA-p02 4</td>
<td>CA-p03 4</td>
</tr>
<tr>
<td>9</td>
<td>20</td>
<td>7</td>
<td>0.0042</td>
<td>CA-p02 5</td>
<td>CA-p03 5</td>
</tr>
<tr>
<td>10</td>
<td>28</td>
<td>14</td>
<td>0.0503</td>
<td>CA/2-p02 7</td>
<td>CA/2-p03 7</td>
</tr>
<tr>
<td>11</td>
<td>28</td>
<td>15</td>
<td>0.1028</td>
<td>CA/2-p02 9</td>
<td>CA/2-p03 9</td>
</tr>
<tr>
<td>12</td>
<td>28</td>
<td>17</td>
<td>0.3164</td>
<td>CA/2-p02 10</td>
<td>CA/2-p03 10</td>
</tr>
<tr>
<td>13</td>
<td>30</td>
<td>16</td>
<td>0.0895</td>
<td>CA/2-p02 10</td>
<td>CA/2-p03 10</td>
</tr>
<tr>
<td>14</td>
<td>25</td>
<td>16</td>
<td>0.4661</td>
<td>CA/2-p02 8</td>
<td>CA/2-p03 8</td>
</tr>
<tr>
<td>15</td>
<td>20</td>
<td>13</td>
<td>0.5211</td>
<td>CA/2-p02 6</td>
<td>CA/2-p03 8</td>
</tr>
<tr>
<td>16</td>
<td>10</td>
<td>6</td>
<td>0.4406</td>
<td>CA/2-p02 5</td>
<td>CA/2-p03 5</td>
</tr>
</tbody>
</table>

- **Columns in order are:** time \( t \) to diagnosis in months; the number of cases (triplets) with measurement taken between \( t \) and \( t + 6 \) months prior to diagnosis; the number of errors when classifying triplets with CA125 alone and associated \( p \)-value; the minimal number of errors when classifying with \( w \log C + v \log I(p) \); most often used peak (see Table II); the ratio of weights \( v/w \) used in the rule; the main \( p \)-value for overall significance of this result; the conditional \( p \)-value for significance of non-CA125 component. Bold areas show significance at \( p \)-value <0.05 for CA125 and main \( p \)-values beyond 9 months.

### Notes:
- The ‘rule’ columns show the best rules selected: CA in these columns means log \( C \); p02 means log \( I(2) \); p03 means log \( I(3) \).
- The number of errors is shown for the best rule with the associated main \( p \)-value and conditional \( p \)-value.
- Conditional \( p \)-value shows the improvement achieved by adding information from peak 2 or 3 beyond 9 months. P-values were multiplied by adjustment coefficients \( 4\zeta(2)=6.579 \) for peak 2 and \( 9\zeta(2)=14.804 \) for peak 3.
The purpose of this study was to assess if low mass serum polypeptides carry information to aid in the early diagnosis of ovarian cancer. We showed that two MALDI peaks (identified as CTAPIII and inferred as PF4, respectively), when combined with serum CA125, provided significantly earlier detection of cancer. CA125 alone gave significant prediction up to 9 months.

**Discussion**

The purpose of this study was to assess if low mass serum polypeptides carry information to aid in the early diagnosis of ovarian cancer. We showed that two MALDI peaks (identified as CTAPIII and inferred as PF4, respectively), when combined with serum CA125, provided significantly earlier detection of cancer. CA125 alone gave significant prediction up to 9 months.
prior to diagnosis, similar to a previous report (36), with MS peak information not adding significantly to this. At greater than 9 months prior to diagnosis, CA125 performance was significantly complemented by adding MS peak information, with confident detection up to 15 months using all peaks and 15 and 11 months using peaks 2 and 3, respectively.

We wanted to confirm the altered expression of these peaks using orthogonal assay methods. Peak 2, speculatively identified as PF4, was assayed using a commercial ELISA. There was no significant difference between cases and controls when all samples were considered, but levels of PF4 were lower in distant versus proximal cases, as well as there being a trend for lower PF4 in cases versus controls at the distant time points, and in women later diagnosed with stage IV cancer. Whilst not confirmatory from a statistical standpoint, these data give some support for reduced serum PF4 levels early in cancer development. For CTAPIII, there were no suitable immunoassays available which would be specific for CTAPIII without also recognising the nine other processed products of platelet basic protein with their overlapping sequences *i.e.* TC-2, CTAPIII(1-81), beta-thromboglobulin, neutrophil-activating peptide 2 (NAP2)(74), NAP2(73), NAP2, TC-1, NAP2(1-66) and NAP2(1-63).

CTAPIII is a chemokine released into the circulation from platelet alpha granules. It is known to stimulate DNA synthesis, mitosis, glycolysis, intracellular cAMP accumulation, prostaglandin E2 secretion and synthesis of hyaluronic acid and sulfated glycosaminoglycan in target tissues. PF4 is also released from platelet alpha granules and possibly leukocytes and has a major role in neutralising the anticoagulant effect of heparin through binding. It is also chemotactic for neutrophils and monocytes and inhibits endothelial and activated T-cell proliferation. From our data, it would appear that the secretion of both chemokines is suppressed in the early stages of ovarian cancer possibly through a host response to tumour development. This suppression may support tumour growth through the modulation of inflammatory and immunogenic processes, coagulation and angiogenesis, and as such may not be specific to ovarian cancer (see below).

Perhaps a weakness of this study is the fact that CA125 was used as part of the diagnostic procedure making interpretation of results for a combined biomarker panel difficult. However, limiting the analysis to only those cases which were screen negative and who developed symptomatic disease would greatly diminish the value of this unique preclinical sample set and the impact of the study. Although the median age was significantly different between the sample groups (*p*=0.01), it has no clinical significance in this post-menopausal population. It is known that HRT use is a risk factor for ovarian cancer (37) and also that HRT use can have a profound effect on the serum proteome (38). However, there was no significant difference in HRT use between the case and control groups used here (Table I) and indeed no correlation between HRT use (or age) and PF4 level (measured by ELISA). We thus conclude that HRT use is not a confounding factor in this study. The other significant difference between the groups was in the use of the oral contraceptive pill (*p*=0.005), which was higher in the control group. This may be expected as the pill is known to confer long-term protection against ovarian cancer (39).

It has been proposed that tumour-specific exopeptidases may generate surrogate peptide markers of cancer *ex vivo* during coagulation (28). Despite this, we recently showed that such peptides do not make useful biomarkers for ovarian cancer diagnosis (32). This result seems at odds with the findings of the present study. However, it is important to note that both CTAPIII and PF4 are released into the circulation in their processed forms and appear not to arise or be subject to proteolysis during serum preparation: we have found no evidence that smaller fragments of these proteins are generated (27) that may explain their reduced levels in pre-diagnosis sera. We therefore hypothesise that these proteins are altered in the earlier stages of tumour development but do not maintain differential expression close to diagnosis. This may explain why we failed to identify changes in these proteins in serum samples from clinically diagnosed cases and controls using the same profiling strategy (32, 40). Here no models incorporating peak intensities were accurately able to discriminate cases from benign or healthy controls, and peak information failed to add to the performance of CA125.

Although a diagnostic test for ovarian cancer based on SELDI-TOF MS assays, includes CTAPIII as an up-regulated protein (8, 41), we have failed to reproduce this using the same strategy (unpublished data). Notably, both CTAPIII and PF4 have been identified as putative biomarkers down-regulated in serum samples from acute lymphoblastic leukaemia cases where they were similarly identified by off-line LC fractionation and ESI-MS/MS following SELDI-TOF MS profiling (34). PF4 was also identified from MALDI-TOF MS serum profiling experiments as a putative biomarker of pancreatic cancer and was shown also to be down-regulated in the samples from cases (35). Although this data implies the poor specificity of these proteins in detecting ovarian cancer, it does support their possible negative roles in cancer progression.

In conclusion, our discovery study shows that the period of significant discrimination in advance of diagnosis can be extended from 9 to 15 months if CA125 is combined with certain MALDI-MS peaks which we identify as the chemokine CTAPIII and putatively as the chemokine PF4. This data supports a link between platelet function and tumour development in the early stages of ovarian cancer. Further work will be required to validate these findings to assess the potential of these markers for early ovarian cancer detection.
Conflict of Interest Statement

IJ has a consultancy arrangement with Becton Dickinson in the field of tumour markers and ovarian cancer but not involving work directly related to this study. UM has a financial interest through UCL Business and Abcodia Ltd in the third party exploitation of clinical trials biobanks developed through research at UCL. This does not involve any interests directly related to this work. The other Authors have no conflicts of interest to declare.

Acknowledgements

Research was undertaken within the Women’s Health Theme of the UCLH/UCL Comprehensive Biomedical Research Centre, the Computer Learning Research Centre, Royal Holloway, University of London and the University of Reading. UCLH/UCL Comprehensive Biomedical Research Centre received a proportion of funding from the Department of Health’s NIHR Biomedical Research Centres funding scheme. The work was supported by MRC grant G0301107 and EPSRC grants EP/E00053/1 and EP/F002998/1.

References


Received August 2, 2011
Revised September 29, 2011
Accepted October 3, 2011