Abstract. Background: The incidence of prostate cancer (PCa) has increased in recent years due to the aging of the population and increased testing; however, mortality rates have remained largely unchanged. Studies have shown deficiencies in predicting patient outcome for both of the major PCa diagnostic tools, namely prostate specific antigen (PSA) and transrectal ultrasound-guided biopsy. Therefore, serum biomarkers are needed that accurately predict prognosis of PCa (indolent vs. aggressive) and can thus inform clinical management. Aim: This study uses surface enhanced laser desorption/ionization time of flight mass spectrometry (SELDI-TOF-MS) mass spectrometry analysis to identify differential serum protein expression between PCa patients with indolent vs. aggressive disease categorized by Gleason grade and biochemical recurrence. Materials and Methods: A total of 99 serum samples were selected for analysis. According to Gleason score, indolent (45 samples) and aggressive (54) forms of PCa were compared using univariate analysis. The same samples were then separated into groups of different recurrence status (10 metastatic, 15 biochemical recurrences and 70 non-recurrences) and subjected to univariate analysis in the same way. The data from Gleason score and recurrence groups were then analysed using multivariate statistical analysis to improve PCa biomarker classification. Results: The comparison between serum protein spectra from indolent and aggressive samples resulted in the identification of twenty-six differentially expressed protein peaks (p<0.05), of which twenty proteins were found with 99% confidence. A total of 18 differentially expressed proteins (p<0.05) were found to distinguish between recurrence groups; three of these were robust with p<0.01. Sensitivity and specificity within the Gleason score group was 73.3% and 60% respectively and for the recurrence group 70% and 62.5%. Conclusion: SELDI-TOF-MS technology has facilitated the discovery of prognostic biomarkers in serum that can successfully discriminate aggressive from indolent PCa and also differentiate between recurrence groups.

Prostate cancer (PCa) is the second most common cause of cancer in men in the Western world (1), resulting in 10,000 UK deaths per year (2). However, it is an unusual cancer in that a far more indolent form of the disease also exists in the majority of men aged over 80 years (3). At present, raised prostate-specific antigen (PSA) levels and a suspicious digital rectal examination (DRE) are used by clinicians to identify patients who require a transrectal ultrasound (TRUS)-guided biopsy of their prostate gland to look for PCa (4). However, PSA is neither particularly sensitive (5) nor specific (6). One shortcoming of PSA is that increases are also found in non-cancerous diseases such as benign prostatic hyperplasia (BPH), chronic inflammation or infection (7-9). PSA is also found in women (10). DRE is very subjective and even TRUS-guided biopsy is imperfect (11). This means that patients with aggressive forms of PCa may be missed and those with indolent disease may be subjected to unnecessary radical therapy and its associated morbidity. Therefore, new biomarkers that can diagnose PCa
with a higher sensitivity and specificity, and better predict prognosis to help guide management are greatly needed. Current research is therefore focused on using proteomic techniques to detect new biomarkers with higher specificity and sensitivity (12). Indeed, it is now moving towards using ‘bio-signatures’ made up of many biomarkers to test for disease, facilitated by novel proteomic technologies such as high-throughput mass spectrometry. One version of this technology is known as surface-enhanced laser desorption time-of-flight mass spectrometry (SELDI-TOF-MS) (13). SELDI-TOF-MS facilitates use of protein chip arrays with different binding affinities such as: metal affinity, IMAC-Cu; hydrophobic, WCX2, C16/H4 and weak cation exchange (CM10). The principle of this method is to extract proteins or peptides and apply them to surfaces enhanced with particular binding chemistries. These binding chemistries act to selectively bind proteins within the sample to form specific subsets. Unbound proteins are washed away and captured proteins are then ionized with SELDI and then measured by TOF-MS. Thousands of protein peaks are generated and displayed as spectra, map, or gel view, which can then be used to distinguish sample differences (14, 15). A computer algorithm then analyses and selects ‘discriminatory’ peaks (16). SELDI-TOF-MS has been proven to have a potential role in biomarker discovery and to be beneficial for early diagnosis, appropriate prognosis, effective therapy and monitoring of cancer treatment (17). Many studies have been published on PCa using SELDI-TOF-MS (18-25). However, only a limited number of studies have investigated PCa progression (26). This study uses SELDI-TOF-MS to identify differentially expressed serum proteins between PCa patients with indolent (Gleason score <7) and aggressive disease (Gleason score ≥7), and between patients with biochemical (BR) or metastatic disease (MR) recurrence and without disease recurrence (NR). Identification of progression-associated biomarkers would help clinicians individualise treatment at an early stage, which is particularly important in patients with a more aggressive form of the disease as there are no treatment options in the latter stages. Furthermore, increased knowledge regarding the course of disease may also limit unnecessary biopsy and surgical procedures which all have their associated risks. To our knowledge, this study represents the first to research candidate biomarkers from indolent and aggressive forms of prostate cancer using SELDI-TOF-MS technology.

Materials and Methods

Serum samples and ethical approval. PCa patients (n=99) were recruited at Queen Alexandra (QA) Hospital, Portsmouth between 2005 and 2008. Only samples from patients retaining their prostate at the time of sampling and with known Gleason score were used. Voluntary written informed consent for the research project was given by all participants and ethical approval for the collection of surplus serum was obtained (Approval number, 03/10/095). Disease was classified according to the histopathology-based Gleason score, with a Gleason score of <7 classed as indolent, and ≥7 classed as aggressive. This cut-off was selected as men with a Gleason score of 7 or above had a greatly increased risk of dying from the disease according to 20-year follow-up data (27). Ninety-nine serum samples from 45 indolent and 55 aggressive cases were used in the study. Samples were taken in a standardised manner using serum separator tubes (SST) for routine PSA testing. SST tubes contain a polymer gel and a clot activator (covering the inner wall of the tube). The sample was incubated for 30 minutes to allow the blood clot to form. It was then centrifuged at 1300-2000xg in 20-25°C for 10 minutes. The supernatant (serum) was obtained by aspiration and placed into 5 ml labelled tube and stored at 4°C. Following routine PSA tests, surplus material was stored for use in this trial. Upon collection, samples were aliquoted into 5x1.7 ml microcentrifuge tubes (50 μl from each patient sample) to minimise freeze/thaw during the study, and then stored at -20°C.

Preparation of serum samples. Pre-analytical parameters were optimised using pooled samples to determine the optimum chip chemistry, buffer and sinapinic acid (SPA) concentration as well as SELDI analysis conditions. This showed that the CM10 chip (weak cation exchange) with 50 mM ammonium acetate buffer pH 4.5 (binding buffer) and 50% SPA in acetonitrile (ACN)/0.5% trifluoroacetic acid (TFA) were optimal for this study (data not shown). Pooling an equal amount of serum from each tumour type (indolent and aggressive) served as a quality control (QC) for assessing the reproducibility and chip variability of the study. Serum samples were defrosted on ice for 90 minutes prior to use and then diluted with a denaturing buffer (U9, 9M Urea, 2% CHAPS, 50mM Tris-HCl pH 9) to generate a 1:10 dilution before being incubated on ice for 30 minutes. CM10 arrays were prepared using a BioMeK 3000 (Beckman Coulter) liquid handling robot. Arrays were pre-wetted with 5 μl of binding buffer for 5 minutes at room temperature, before two additions of 200 μl of binding buffer for 10 minutes at room temperature with shaking on a Micomix 5 platform shaker (DPC; 20, amplitude 7). A total of 10 μl of U9-treated serum was then added to 90 μl of binding buffer and incubated for 60 minutes at room temperature with shaking. The arrays were then washed three times with 200 μl of binding buffer for 10 minutes each, followed by three HEPES (5 mM HEPES, pH 7) rinses to eliminate any unbound proteins or buffer salts which could interfere with array analysis. Prior to SELDI-TOF-MS analysis, the arrays were air dried for 45 minutes and then treated with two additions of 1 μl of 50% SPA in 50% TFA. All samples were applied to the chips in duplicate (to minimise drop out) and a QC sample was included on one spot of each chip from all arrays prepared.

SELDI-TOF-MS protein profiling. The Human Genetics Division proteomics laboratory at the University of Southampton provided the Enterprise 4011 SELDI-TOF-MS platform for PCa serum sample analysis. Protein Chip Data Manager v3.0.7 Software (PCDM) was used for analysis. A mass range between 0-10 kDa was examined with a focus mass of 5 kDa and matrix attenuation at 1 kDa, whereas a mass range above 10 kDa was studied with a 19 kDa focus mass and 5kDa matrix attenuation. External calibration of the SELDI-TOF-MS was carried out using a calibration kit of protein standards comprised of: recombinant hirudin (6.96 kDa), equine cytochrome c (12.23 k Da), equine...
myoglobin (16.95 kDa), and carbonic anhydrase (29.0 kDa). Mass accuracy was found to be approximately 0.02% of actual mass value. Noise was eliminated using the noise definition setting which removes chemical noise resulting from the energy absorbing matrix in the low mass range. Therefore, only the area above matrix attenuation (1 kDa) was usable. Clusters were composed of peaks with a minimum signal-to-noise (S/N) ratio of 5 and valley depth of at least 5, using peaks which were present in at least 20% of the relevant spectra. The average mass of clusters was then used to identify unlabelled spectra, and therefore, each spectrum showed a peak intensity value. The mass window for each cluster was identified as 0.3% of the peak mass for low mass (0-10 kDa) and high mass range (>10 kDa) of the spectra optimised. The validity of peaks used as candidate biomarkers was confirmed by further evaluation and relabelling prior to final statistical analysis. Automatic detection was used to identify qualified mass peaks with a signal-to-noise ratio of more than five and a mass value range of 1-200 kDa. Further peak clusters were added after applying a second peak selection with (S/N of >2, within 0.3% mass window).

Data processing, normalisation, peak detection, feature selection and statistical analysis. Baselines were subtracted from the spectra which were then normalised within the mass range of 1-200 kDa. The normalisation factor was used to determine the outlier spectra showing possible variation caused by sample processing. Spectra with a normalisation factor of two-fold higher than the average were deleted or replaced with duplicate spectra. Univariate analysis was then performed using Mann-Whitney U-tests to compare peak intensity between two groups. Ninety-nine serum samples from PCa patients with Gleason score and recurrence data were compared using univariate analysis. The same samples were then separated into recurrence status groups (10 MR, 15 BR and 70 NR) and subjected to univariate analysis. For multivariate analysis, the mass (m/z) and intensity values for high and low mass ranges were combined and exported to build the tree algorithm. Multivariate statistical analysis was then used to classify the samples and evaluate the sensitivity and specificity using a tree classification algorithm. Multivariate analysis was performed using 69 (39 aggressive, 30 indolent) randomly assigned serum samples as a training set and 30 (15 indolent, 15 aggressive) as a test set. Multivariate analysis of recurrence was performed using 67 (10 BR, 7 MR, 50 NR) and 28 (5 BR, 3 MR, 20 NR) serum samples as training and test sets, respectively. Four samples were excluded from the recurrence analysis due to lack of recurrence information. To assess variation, the coefficient of variation (CV) of pooled QC samples was calculated using the formula:

\[ CV = \sqrt{CV_1^2 + CV_2^2 + CV_n^2}/n \]

where: \( n \) = the number of peaks in the spectra.

Results

Serum candidate biomarkers categorised by Gleason score. Using the SELDI-TOF-MS software to study all spectra, 43 and 50 clusters of peaks were generated in the low mass and high mass ranges, respectively. The intensity of these clusters was compared between indolent and aggressive groups and yielded 26 statistically significant (\( p<0.05 \)) differentially expressed peaks. These differentially expressed peaks included 18 underexpressed and 8 over-expressed biomarkers in the aggressive group relative to the indolent cohort (Table I). Examples of the differentially expressed protein peaks are shown in Figure 1. Interestingly, all of the underexpressed biomarkers occurred within the high mass range and all of the over-expressed biomarkers occurred within the low mass range. Twenty-two biomarkers were the most plausible ion peaks (\( p>0.01 \)) with 99% confidence between the two groups. The other four were significant at (\( p<0.05 \)). Among protein peaks, 8 protein peaks were up regulated (↑) and 18 were down regulated (↓) in serum samples of patients with aggressive disease.

<table>
<thead>
<tr>
<th>Mean m/z</th>
<th>p-Value</th>
<th>Mean intensity</th>
<th>Expression change in aggressive disease</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Indolent</td>
</tr>
<tr>
<td>4644.07</td>
<td>0.00056</td>
<td>1.53</td>
<td>2.20</td>
</tr>
<tr>
<td>9298.54</td>
<td>0.00098</td>
<td>7.79</td>
<td>10.27</td>
</tr>
<tr>
<td>2746.60</td>
<td>0.00295</td>
<td>1.66</td>
<td>2.99</td>
</tr>
<tr>
<td>7770.51</td>
<td>0.00563</td>
<td>22.50</td>
<td>28.12</td>
</tr>
<tr>
<td>3880.35</td>
<td>0.00613</td>
<td>2.77</td>
<td>3.47</td>
</tr>
<tr>
<td>4615.73</td>
<td>0.00913</td>
<td>1.30</td>
<td>2.25</td>
</tr>
<tr>
<td>3888.01</td>
<td>0.02197</td>
<td>2.44</td>
<td>2.99</td>
</tr>
<tr>
<td>4396.25</td>
<td>0.03441</td>
<td>1.40</td>
<td>2.08</td>
</tr>
<tr>
<td>4451.27</td>
<td>0.00019</td>
<td>2.82</td>
<td>2.27</td>
</tr>
<tr>
<td>100608.45</td>
<td>0.00040</td>
<td>1.41</td>
<td>0.77</td>
</tr>
<tr>
<td>28936.34</td>
<td>0.00052</td>
<td>4.10</td>
<td>2.45</td>
</tr>
<tr>
<td>22226.88</td>
<td>0.00068</td>
<td>2.95</td>
<td>2.30</td>
</tr>
<tr>
<td>14049.68</td>
<td>0.00073</td>
<td>16.82</td>
<td>10.97</td>
</tr>
<tr>
<td>28101.22</td>
<td>0.00077</td>
<td>16.44</td>
<td>9.74</td>
</tr>
<tr>
<td>89702.31</td>
<td>0.00087</td>
<td>1.28</td>
<td>0.68</td>
</tr>
<tr>
<td>133986.69</td>
<td>0.00094</td>
<td>12.96</td>
<td>8.30</td>
</tr>
<tr>
<td>111440.66</td>
<td>0.00139</td>
<td>0.36</td>
<td>0.27</td>
</tr>
<tr>
<td>168304.09</td>
<td>0.00306</td>
<td>0.09</td>
<td>0.05</td>
</tr>
<tr>
<td>95129.31</td>
<td>0.00486</td>
<td>0.70</td>
<td>0.42</td>
</tr>
<tr>
<td>178854.19</td>
<td>0.00552</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td>141468.14</td>
<td>0.00552</td>
<td>1.90</td>
<td>1.54</td>
</tr>
<tr>
<td>157657.89</td>
<td>0.00576</td>
<td>0.14</td>
<td>0.08</td>
</tr>
<tr>
<td>118075.77</td>
<td>0.00756</td>
<td>0.10</td>
<td>0.07</td>
</tr>
<tr>
<td>33315.15</td>
<td>0.00788</td>
<td>15.05</td>
<td>11.50</td>
</tr>
<tr>
<td>79198.54</td>
<td>0.01155</td>
<td>3.71</td>
<td>2.91</td>
</tr>
<tr>
<td>66685.96</td>
<td>0.01575</td>
<td>99.04</td>
<td>77.13</td>
</tr>
</tbody>
</table>

Table I. SELDI protein peaks in serum discriminating the indolent group from the aggressive group. The intensity of clusters was compared between indolent and aggressive groups and yielded 26 statistically significant differentially expressed peaks. Twenty-two biomarkers were the most plausible ion peaks (\( p>0.01 \)) with 99% confidence between the two groups. The other four were significant at (\( p<0.05 \)). Among protein peaks, 8 protein peaks were up regulated (↑) and 18 were down regulated (↓) in serum samples of patients with aggressive disease.
Figure 1. Representative spectra of SELDI TOF-MS protein peaks: a: 4643 Da; b: 4397 Da; c: 9298 Da; d: 22226 Da and e: 44583 Da. Protein peaks a-c were overexpressed in aggressive disease compared to indolent disease, while those shown in d and e were overexpressed in the indolent group compared to the aggressive group. The X-axis shows the molecular weight; the Y-axis shows the relative peak intensity.
with differentially expressed proteins identified via the univariate analysis, although 44583 Da was not. However, a slightly smaller peak of 44512 Da was identified via the univariate analysis, hence this is likely to be the same protein. The optimal tree was then used to assign patients from the test set into groups. A receiver operator characteristics (ROC) curve was used to evaluate the tree produced from the training set. The area under the curve (AUC) for the ROC was calculated as 0.897. The tree was then applied to the independent test set and the sensitivity and specificity calculated as 73.3% and 60% respectively (Table II).

Table II. Validation of classification tree algorithm with 30 test samples. The classification tree algorithm was generated using CART software (Classification and Regression Trees). The nodes of the tree were then applied to the independent test set (15 aggressive vs. 15 indolent). The sensitivity and specificity were calculated as 73.3% and 60% respectively.

<table>
<thead>
<tr>
<th>Actual cases</th>
<th>Total cases</th>
<th>Percentage correct</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Aggressive</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N=17</td>
</tr>
<tr>
<td>Aggressive</td>
<td>15</td>
<td>73.33</td>
<td>11</td>
</tr>
<tr>
<td>Indolent</td>
<td>15</td>
<td>60.00</td>
<td>6</td>
</tr>
</tbody>
</table>

Table III. Validation of classification tree algorithm with 28 test samples. The classification tree algorithm was generated using CART software (Classification and Regression Trees). The nodes of the tree were then applied to the independent test set (20 NR vs. recurrence groups BR and MR). The sensitivity and specificity were calculated as 70% and 62.5% respectively.

<table>
<thead>
<tr>
<th>Actual cases</th>
<th>Total cases</th>
<th>Percentage correct</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Aggressive</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N=17</td>
</tr>
<tr>
<td>NR</td>
<td>20</td>
<td>70.00</td>
<td>14</td>
</tr>
<tr>
<td>BR + MR</td>
<td>8</td>
<td>62.50</td>
<td>3</td>
</tr>
</tbody>
</table>

NR: No recurrence; BR biochemical recurrence; MR: metastatic recurrence.

Serum candidate biomarkers categorised by recurrence status. Using the SELDI TOF-MS software to study all spectra, 43 and 50 peak clusters were generated in the low mass and high mass ranges respectively. Univariate analysis across the three groups (NR, BR and MR) was performed and yielded twelve putative biomarkers with p-values <0.05. Three of these were significant with 99% confidence (p≤0.01; Table II).

Figure 3a, b). These biomarkers were able to differentiate between NR, BR and metastatic groups. Four biomarkers (6450 Da, 6650 Da, 3079 Da and 8941 Da) showed an increase in expression associated with BR, and by a decrease associated with metastatic disease in comparison to patients with NR disease (Figure 3a). Conversely, six proteins (7190 Da, 7482 Da, 8352 Da, 8135 Da, 4061 Da and 7929 Da) showed down-regulation with BR but a similar level of expression in metastatic disease compared to NR disease (Figure 3b). One protein peak (12869 Da) had a similar level of expression in NR and BR, with an increase associated only with metastatic disease (Figure 3c). Another protein (2747 Da) had a similar expression level in BR and metastatic disease with a lower expression level evident in NR disease (Figure 3c). Interestingly, the latter protein also showed an increased expression in aggressive disease based on Gleason score (Table III). Comparisons of NR and BR groups identified five putative biomarkers (6856 Da, 3888 Da, 3152 Da, 4466 Da and 16254 Da) that differed significantly (Figure 3d). These markers were not identified via the three group analysis above. Four of these (6856 Da, 3888 Da, 3152 Da and 4466 Da) showed an increased expression in the BR group
compared to the NR group, with only one (16254 Da) showing reduced expression. Furthermore, protein peak 3888 Da was also identified as being up-regulated in aggressive disease based on Gleason score. In addition to the markers identified from comparisons of all three recurrence groups, another marker was found to be significantly different between BR and MR groups. A protein peak of 118180 Da was found to show increased expression in metastatic disease compared to BR (Figure 3e).

From the peak clusters, 67 serum samples from the training set and 28 from the test set were evaluated by multivariate analysis. Due to the limited numbers of samples from patients with recurrence, BR and MR groups were combined for multivariate analysis. A tree with 1 node (3314 Da) resulted from the multivariate analysis (Figure 4) with a relative ‘cost’ of 0.45. The calculated AUC for the ROC was 0.771 and when validated using the test set, yielded a sensitivity and specificity of 70.0% and 62.5% respectively (Table III). However, this marker (3314 Da) was
not identified as being significantly differentially expressed by univariate analysis. Examples of the differentially expressed protein peaks are shown in Figure 5. An increased level of expression of 3314 Da was associated with recurrent disease (BR or MR; Figure 5c).

Discussion

The use of PSA as a diagnostic marker for PCa is not ideal due to its lack of specificity and sensitivity (5, 6). Many studies have tried to refine this by assessing PSA characteristics such as PSA density, PSA velocity, PSA doubling-time and age-specific PSA ranges (28-31), however these modifications have still failed to overcome the shortcomings of current PSA testing (5, 32). There is therefore a need to find new and better biomarkers for PCa that can lead to an earlier diagnosis and differentiate between indolent and aggressive forms of the disease (33).

Serum proteomic profiling has been vetted as a new diagnostic approach in PCa research. These profiles may be used to detect differences between malignant and benign disease and to differentiate between different subtypes of the same disease, such as indolent and aggressive forms of PCa (34). Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) was the first technique developed for the identification of protein biomarkers from complex clinical samples and, until recently, gave the best snapshot of protein samples derived from cells (35, 36). However, 2D-PAGE is time-consuming, labour intensive and, more importantly, is unable to detect proteins smaller than 10,000 Da (37).

Figure 5. Representative spectra of the peaks with a molecular mass of 6650 Da, 7190 Da, and 3314 Da obtained from SELDI-TOF-MS. a: A peak of 6650 Da with increased expression in BR and a decreased expression in MR and NR groups. b: A protein with mass 7190 Da was of low abundance in BR and high abundance in MR and NR groups. c: A protein of mass 3314 derived from multivariate analysis was found to be up-regulated in NR and down-regulated in recurrence groups. The X-axis shows the molecular weight; the Y-axis shows the relative peak intensity.
SELDI-TOF-MS overcomes this problem and has enabled the detection of this potentially important group of proteins (22, 38-41). The main advantages of the SELDI approach are its high throughput (42) ability to analyse proteins in body fluids directly on protein chip arrays without the necessity of long pre-analytical handling processes (43). Moreover, SELDI-TOF-MS technology can be applied to samples of low volume (1 μl) obtained from serum (44). It has been demonstrated that SELDI-TOF-MS has the potential role of obtaining high sensitivity, specificity and resolution for new biomarkers that are detected from different biological samples (45, 46). The analysis of samples using SELDI-TOF-MS may be easier than with other technologies as it does not require pathophysiological knowledge of the disease prior to analysis on the protein chip array (47). SELDI-TOF-MS has been utilised in many applications other than seeking new protein profiles, such as studies of protein–protein interactions (48, 49), protein–DNA interactions (50) and post-transitional modifications (51). However, SELDI faces hurdles that can limit its role in biomarker discovery. One of these obstacles is the reproducibility of protein chip arrays for samples that are analysed (52). The CV from array to array can influence the intensities of the peaks by 10–40% (53). To overcome this variation, researchers contribute to the development of new methods to ensure the consistency of results. Research towards obtaining reproducible and standardised results is required if protein profiling derived from SELDI-TOF-MS is to become a diagnostic and prognostic tool (54). Furthermore, the high cost of SELDI-TOF-MS may be prohibitive in widespread research (47).

In our study, the CM10 chip was adopted to analyse 99 serum samples from PCa patients. With masses between 2500 and 200000 Da, twenty-six differentially expressed proteins \(p < 0.05\) were generated using Gleason score to classify patients as having aggressive or indolent disease: 18 biomarkers were underepressed, and 8 were overexpressed in aggressive compared to indolent disease groups. From multivariate statistical data, a three-node tree was produced with an acceptable relative cost. The sensitivity and specificity of the model was 73.3% and 60% respectively when applied to the smaller test set. In terms of recurrence, specificity of the model was 73.3% and 60% respectively with an acceptable relative cost. The sensitivity and multivariate statistical data, a three-node tree was produced in aggressive compared to indolent disease groups. From classify patients as having aggressive or indolent disease: 18 proteins (2500 and 200000 Da, twenty-six differentially expressed serum samples from PCa patients. With masses between TOF-MS may be prohibitive in widespread research (47). These studies are not comparable with our own research because animal rather than human samples were used (60), or advanced stage disease was examined (19, 61). Our own study was categorised by Gleason score \(<7\) or \(\geq 7\) and recurrence status (NR, BR and MR).

The training and test sets require high numbers of samples to achieve reliable results. Adam et al. (18) used a training set composed of 167 PCa, 77 BPH, 82 healthy control cases and a test set with 30 PCA, 15 BPH and 15 healthy controls. The sensitivity and specificity were 83% and 97% respectively for PCA vs. healthy controls. Qu et al. (57) used a similar number of samples for their training and test sets (18). They identified diagnostic PCA models that were able to predict PCa with a sensitivity of a 97% and...
a specificity of 97%. In our study, the sample size for training and test set was limited and therefore, a slightly lower sensitivity and specificity was achieved. Additionally, utilising multiple chip surfaces can result in a greater number of biomarkers and a more robust model. A study by Banez et al. (58) analysed samples using IMAC and CM10 chips. They analysed 44 PCa, 30 HC as a training set and 62 PCa and 26 healthy controls as test set. The sensitivity produced from CM10 chip was 63% and specificity was 77%. The sensitivity produced from IMAC chip was 34% and the specificity was 62%. However, 85% sensitivity and specificity was produced when IMAC and CM10 data were combined. The sensitivity and specificity for the model produced from the CM10 chip was similar to that produced by our study. This could be due to both their study and ours having similar training sample sizes.

In conclusion, this study represents the first investigation of candidate biomarkers differentiating indolent and aggressive forms of PCa based on SELDI-TOF-MS technology. The first, proof-of-concept stage of our research has alluded to proteomic differences which now require validation of the resultant biomarkers by immuno approaches.

Acknowledgements

We would like to thank Dr Sultan Al-Sadiri and our colleagues Professor Ian Cree, Mr Simon Holmes and Mrs Vanessa Basetter for their continued support and Dr Alan Cooper for his advice. We would like to thank Madeline Fiji Schuck-Ranera and Mary Jane Anomuevo for their assistance. We would also like to thank the funding bodies for their commitment to our work: the Saudi Arabian Government (Ministry of Higher Education), and Wessex Medical Research.

References


