Research paper

One-step immunostrip test for the simultaneous detection of free and total prostate specific antigen in serum

César Fernández-Sánchez a,1, Calum J. McNeil a,*, Keith Rawson b, Olle Nilsson c, Hing Y. Leung d, Vincent Gnanapragasam d

a School of Clinical and Laboratory Sciences, The Medical School, University of Newcastle upon Tyne, Framlington Place, Newcastle upon Tyne, NE2 4HH, UK
b Cambridge Life Sciences Ltd., 14 St. Thomas’ Place, Cambridge Business Park, Ely, Cambridgeshire CB7 4EX, UK
c CanAg Diagnostics AB, Majnabbe Terminal, SE-414 55, Gothenburg, Sweden
d School of Surgical and Reproductive Sciences, The Medical School, University of Newcastle upon Tyne, Framlington Place, Newcastle upon Tyne, NE2 4HH, UK

Received 30 March 2005; received in revised form 29 July 2005; accepted 17 August 2005
Available online 24 October 2005

Abstract

The development of a one-step lateral flow immunoassay on a strip format for the rapid and reliable simultaneous detection of serum levels of free and total prostate specific antigen (f-PSA and t-PSA) and estimation of f-PSA to t-PSA ratio (f/t-PSA) is reported. The f/t-PSA ratio has shown to be more specific for the correct diagnosis of prostate cancer than t-PSA alone, especially in the so-called diagnostic grey zone of 4–10 ng/l t-PSA. The performance of the system described relied on non-competitive immunoassay protocols. Herein, f-PSA and t-PSA were sandwiched between anti-f-PSA and anti-t-PSA monoclonal antibodies immobilised on the strip and a colloidal gold anti-PSA antibody tracer. In the presence of PSA in the sample, the tracer accumulated on the strip results in the appearance of specific pink colour lines. The colour intensity of these lines was found to be directly proportional to the PSA concentration and a semi-quantitative estimation could be carried out visually. Quantitative analysis was also possible by densitometry. Using PSA standards prepared in female serum, the strip could be calibrated up to a concentration of 60 ng/l for both PSA species, with an assay time of less than 20 min. The estimated detection limit was 1 ng/l in all cases. The immunostrip showed good storage stability for at least 2 months and the reproducibility was always between 12% and 17%.

Fifty-one male serum samples were analysed with the strip and results compared with values obtained by two different commercial immunoassays taken as reference methods. The study yielded acceptable correlation and agreement. An estimation of the sensitivity and specificity demonstrated the strip validity as a potential front-line device for the early detection of prostate cancer and differentiation of benign prostatic anomalies.

Small plastic cartridges incorporating the immunostrip and a small blister containing washing solution that helps remove unbound species from the strip were envisaged in order to avoid false positive readings and decrease background signals, thereby leading to better sensitivity and detection limits.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Lateral flow assay; Immunostrip; Free and total prostate specific antigen; Colloidal gold; Serum

* Corresponding author. Fax: +44 1912227991.
E-mail address: calum.mcneil@ncl.ac.uk (C.J. McNeil).
1 Present address: Centro Nacional de Microelectrónica, CNM-IMB-CSIC, Barcelona, Spain.
0022-1759/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.jim.2005.08.014
1. Introduction

According to the World Health Organization (WHO), prostate cancer (PCa) is the most common malignancy and the second leading cause of cancer mortality among the male population in developed countries (WHO, 2004). PCa arises mostly as a result of age, race, and a family history of the disease but it may also be related to dietary behaviour. Implementation of screening programmes for asymptomatic men at high risk has been considered to help detect prostate cancer cases at an organ-confined stage thus reducing high risk of PCa. Presently, screening for cancer cases at an organ-confined stage is carried out by the measurement of serum prostate specific antigen (PSA) and digital rectal examination. The introduction of the first PSA immunoassay as a clinical test has led to a dramatic decrease in the number of patients with metastatic disease at the time of PCa diagnosis (Han et al., 2004).

Serum PSA is currently the most reliable tumour marker for diagnosis of prostate cancer and disease recurrence after treatment (Polascik et al., 1999). PSA is a 33 kDa single-chain glycoprotein with protease activity synthesised in the prostate gland and secreted in high concentrations into the seminal fluid. Its function is to liquefy the seminal coagulum by proteolysis. A very low proportion escapes into the bloodstream, where it circulates in a number of molecular configurations and complexes. The predominant molecular form is the complex of PSA with α₁-antichymotrypsin (PSA-α₁ AT) (Lilja et al., 1991). Non-complexed “free” PSA (f-PSA) represents a small but variable proportion of the total PSA concentration. Measurement of serum t-PSA content includes all of the above-mentioned immunoreactive forms of PSA (Milford Ward et al., 2001).

Serum t-PSA concentration in healthy males is in the range of 0–4 μg/l and increases in men with prostate cancer (Karazanashvili and Abrahamsson, 2003). However, the measurement of t-PSA alone lacks specificity since other benign prostate anomalies such as prostatitis or benign prostatic hyperplasia (BPH) may induce an increase in its serum content. In clinical practice PSA levels of above 10 μg/l indicate a high risk of significant prostate cancer. Failure in the correct diagnosis of PCa is more acute in the t-PSA range of 4–10 μg/l (the so-called diagnostic grey zone). Within this range, percentage f-PSA measurement (f-PSA to t-PSA ratio, f/t-PSA) has been demonstrated to significantly improve cancer specificity (Catalona et al., 1995). f/t-PSA is usually lower in men with prostate cancer than in those with benign prostatic conditions (Veltri and Miller, 1999). It has been reported that a cut-off value of 0.25 improved diagnosis specificity by avoiding 20% of unnecessary biopsies while 95% of cancers within the diagnostic grey zone were still correctly detected (Catalona et al., 1998).

Since the launch of the first test for the measurement of serum total PSA in 1980 (Kuriyama et al., 1980) numerous assays for f-PSA, PSA-ACT and t-PSA have become commercially available. All of them rely on different immunoassay formats. Significant inter-assay discrepancies were firstly encountered due to the use of polyclonal anti-PSA antibody sera showing significant differences in affinity and specificity against the various serum PSA forms and the lack of standardisation of PSA calibration solutions (Cheli et al., 1998). The introduction of equimolar anti-PSA antibodies showing identical reactivity against f-PSA and PSA-ACT has led to a better agreement among PSA antibodies from different manufacturers (Nilsson et al., 1997). Likewise, the establishment by WHO of two international standards for PSA, that is a mixture of purified PSA-ACT and f-PSA in a 90:10 proportion and a purified f-PSA solution, further decreased inter-assay variability (Rafferty et al., 2000).

Most of the PSA assays currently on the market are formulated for use on specific immunoassay platforms in clinical chemistry laboratories. Moreover, they are expensive, require skilled personnel and are associated with time-consuming sample processing so that results are not instantly available. Therefore, these tests are not ideally suited for the establishment of large screening programmes to be potentially carried out in the urological clinic and general practice environments.

Different promising approaches for rapid and decentralised immunoassay testing include immunosensors and immunotests. Immunosensor technology makes use of immunochemicals as molecular recognition elements to construct self-contained devices that also comprise a physicochemical transducer (Thévenot et al., 2001). Due to the high affinity constants of the antigen–antibody interactions, most of them are single-use systems. In spite of the high selectivity and sensitivity that immunosensors normally exhibit, the lack of stability, limited reproducibility and the necessity for the addition of external reagents of most of the approaches developed so far are probably the major reasons for their limited commercial success. A number of immunosensor designs for the detection of PSA that rely on optical (fluorescence, laser-based) (O’Neill et al., 1995; Wu et al., 2001), electrochemical (amperometric, impedimetric) (Sarkar et al., 2002; Fernández- Sánchez et al., 2004a,b) and piezoelectric transduction...
modes (Lee et al., 2005) have been reported. Within these techniques only the fluorescent approach was developed by a commercial company and assessed in clinical samples. However, to the best of our knowledge this has not been translated into a viable commercial device. With regard to the remaining approaches reported, PSA has been used rather as a model analyte to demonstrate new technology for the development of immunosensors.

Up to now, membrane-strip test devices in either lateral flow or flow through immunoassay formats have provided the major alternatives to centralised immunochemical testing (Bühler-Sekula et al., 2000). Immunostrip performance relies on the migration of samples and reagents along antibody-coated membrane strips where the corresponding affinity interactions take place and can be detected in just a few minutes without the need for any separation or washing steps (Chan et al., 2003; Müller-Bardoff et al., 1999). Most immunostrip tests are based on the use of colloidal gold antibody tracers, which, due to the pink colour of the label, may be detected visually, thus providing rapid on-site qualitative information, i.e. yes/no response, or semi-quantitative information, by comparison with a colour scale card, about the target analyte (O’Keeffe et al., 2003). Also, quantitative analytical data can be collected by measuring colour intensity by densitometry (Leung et al., 2003). This technology has given rise to some very popular commercially available diagnostic tools such as pregnancy or blood glucose tests, and has proved to be cheap, robust and user-friendly.

Qualitative and semi-quantitative t-PSA immunostrips that work with either serum or whole blood samples have found widespread acceptance in the marketplace. However, comparative clinical trials carried out with some of these devices demonstrated a lack of performance uniformity (Jung et al., 1999) and failed to give accurate measurements especially at PSA values close to 4 μg/l (Lein et al., 1996; Madersbacher et al., 1996). Indeed, tests at low PSA concentrations proved to be difficult to read and interpret due to the appearance of faint colours of the specific affinity reaction on the capture area of the strip, which might be masked by poor background signals (Oberpenning et al., 2003).

A prototype of a strip-based immunosensor device was developed by our research group and demonstrated as a feasible point-of-care tool for the separate detection of f-PSA and t-PSA (Fernández-Sánchez et al., 2004a,b). This approach combines the advantageous features of a lateral flow immunostrip with the sensitivity, ease of fabrication and integration with a capacitive-based transducer. However, it has not yet been tested with real samples. Also, although the simultaneous measurement of both f-PSA and t-PSA was envisaged, slight alterations of the current prototype have yet to be carried out in order to facilitate these measurements.

The aim of the present study was to develop a more straightforward approach for the simultaneous detection of both PSA species. A one-step lateral flow immunostrip that permitted the rapid and quantitative determination of f/t-PSA ratio in serum was fabricated and its performance assessed by comparison with results obtained using two commercial immunoassay kits for free and total PSA. The affinity reactions on the strip relied on non-competitive immunoassay formats using monoclonal antibodies with equimolar reactivity against both f-PSA and PSA-ACT forms, and a colloidal gold label. Calibration curves were obtained using f-PSA standards prepared in male serum. Removal of unbound species from the strip by introducing a washing step was desirable in order to decrease background readings and thus improve strip sensitivity and reliability. Furthermore, a cassette that integrated the developed immunostrip with a small storage blister containing washing solution was envisaged (Long et al., 2002). This cartridge also included a mechanism that stopped the sample flow on the strip and released the washing solution from the blister in an automatic fashion. A preliminary approach towards the realization of this device was tested.

2. Materials and methods

2.1. Antibodies

PSA66 monoclonal antibody (mAb) anti-t-PSA, PSA10 mAb anti-t-PSA, PSA19 mAb anti-f-PSA, PSA20 anti-f-PSA and PSA30 mAb anti-f-PSA (IgGs, all developed in mouse) were kindly supplied by CanAg Diagnostics AB (Gothenburg, Sweden). Goat anti-mouse IgG (γ-chain specific) was purchased from Sigma Chemical Co. (Gillingham, Dorset, UK).

All anti-PSA antibodies were supplied in 0.15 M NaCl solutions without any preservatives. They were aliquoted and stored at −20 °C until use. PSA66 mAb detects an exposed epitope shared by PSA and human kallikrein 2. PSA10 mAb detects an epitope exposed in both f-PSA and PSA-ACT, which is specific for PSA. PSA19, PSA20 and PSA30 mAbs detect different hidden epitopes covered by ACT thus specific to free PSA.

PSA66 mAb was labelled with colloidal gold (40 nm particle size; OD at 530 nm was 11.04) by Alchemy Laboratories Ltd. (Dundee, Scotland, UK). The resul-
ting anti-PSA colloidal gold conjugate stock solution was stored at 4 °C and used as received.

2.2. PSA standards and serum samples

An f-PSA standard was kindly supplied by CanAg Diagnostics AB (Gothenburg, Sweden). This was a stock solution prepared in 0.1 M phosphate buffer pH 7.2 containing 0.15 M NaCl and 60 g/l BSA. A PSA-ACT standard was purchased from Scripps Laboratories (San Diego, CA, USA). It was supplied in 0.01 M acetate buffer pH 5.6 containing 0.15 M NaCl and 0.1% sodium azide. Both standards were diluted to a convenient concentration in their corresponding purchase buffer, aliquoted, and stored at −20 °C. Different dilutions of both PSA solutions were prepared and tested several times during the course of this investigation in order to verify their PSA concentration (f-PSA and PSA-ACT) and to check the storage stability of the stock solutions.

Fifty-one male and ten female serum samples were collected with ethical approval and stored immediately on wet ice. The samples were centrifuged within 2 h of collection. The resulting serum specimens were stored at −80 °C until the analyses were carried out.

Female serum samples that tested negative for PSA were spiked with either f-PSA or PSA-ACT using those solutions mentioned above in order to obtain calibrators in the range of 1–100 µg/l.

2.3. Materials and fabrication of immunostrips

Porex® Lateral-Flo™ membrane was purchased from Mupor Ltd. (Alness, Scotland). Unlike Nitrocellulose membranes, commonly used for immunostrip development, this polyethylene-based membrane shows controlled pore size and shape, stable hydrophilic properties, high protein binding under different experimental conditions and rapid capillary rise. GF/DVA and 27Q glass fibre materials and D28 cellulosic paper were purchased from Whatman International Ltd. (Maidstone, Kent, UK). SS903 cotton linter paper was obtained from Schleicher and Schuell UK Ltd. (London, UK).

Porex membranes were cut into strips 25-mm wide by 250-mm long. Three antibody capture lines were deposited on the membrane strip using appropriate dilutions of the antibody solutions to f-PSA and t-PSA, and goat anti-mouse IgG solution, prepared in 0.1 M phosphate buffer pH 7.2 containing 0.15 M NaCl. First, a control line was obtained by carefully dragging a pipette tip containing 100 µl of the anti-mouse IgG dilution along the strip at as even a speed as possible. This was carried out by leaning the tip against a ruler placed at 6 mm from one end of the membrane. Immediately afterwards, the same process was repeated in order to generate test lines for t-PSA and f-PSA. A 5-mm gap was left between two consecutive lines. Subsequently, the membrane was allowed to dry fully for at least 2 h at room temperature and then applied to a backing material. An 8-mm-width strip of D28 absorbent material was also applied at the downstream end of the membrane strip. This material was used as the sink pad. The resulting construct was then cut into 6-mm-width test strips. Next, 7-mm-length 27Q glass fibre pads were impregnated with 15 µl of undiluted colloidal gold anti-PSA antibody conjugate solution and left to dry overnight at RT. The resulting conjugate release pad together with a 5-mm-length SS903 sample addition pad were placed at the opposite end of the sink pad in order to complete the immunotest strip (Fig. 1A). These strips were stored at either 4 °C or RT under desiccated conditions until use.

2.4. Lateral flow immunoassay

The performance of the immunostrip relied on non-competitive assay formats, where PSA species were sandwiched between the primary anti-PSA antibody immobilised on the strip and the secondary anti-PSA antibody-colloidal gold conjugate (Fig. 1B). Some 80 µl of a sample solution were added to the sample pad and allowed to flow through the conjugate pad. By these means the conjugate dissolved and was pulled to flow through Fig 1. (A) PSA immunostrip. All pads overlap the adjacent material by around 2 mm. (B) Non-competitive colloidal gold based immunoassay. 1—f-PSA capture line; 2—t-PSA capture line; 3—control line (anti-mouse IgG); 4—sample pad; 5—conjugate pad; 6—sink pad; 7—backing material; 8—PSA; 9—colloidal gold antibody tracer.
along the membrane together with the sample solution. If PSA in its different forms was present in the sample, it bound on to the corresponding antibody capture lines on the strip together with the antibody-colloidal gold conjugate. As a result, pink colour lines appeared on the capture zones of the strip if the PSA content (f-PSA and/or PSA-ACT) was above the detection limit of the device. Any excess unbound conjugate produced a pink colour on the control line, this being an indication of the technical validity of the test. Finally, the solution reached the sink pad. The overall flowing time was always less than 1 min. After a 15-min incubation period, the reaction was stopped by removing the sample, the conjugate and sink pads. The strip was immediately washed backwards (to the opposite direction of the sample flow) and sample flow was based on previous studies carried out by our research group. Nitrocellulose membranes were tested together with Porex polyethylene-based material. Porex membranes were found to be more robust and did not require a blocking step following antibody coating. In order to prevent conjugate non-specific binding, nitrocellulose membranes do require either this step or the addition of blocking agents to the sample pad in most applications. This normally has a negative effect on their capillary flow properties and the overall performance of the test (Verheijen et al., 1998).

2.5. Immunoassays for method comparison

Male serum samples were also analysed using two commercial PSA immunoassay kits that rely on enzyme labels and two different detection modes. CanAg free PSA and total PSA enzyme immunoassays from CanAg Diagnostics AB (Gothenburg, Sweden) consist of non-competitive solid-phase assay protocols with spectrophotometric detection. The detection limits for t-PSA and f-PSA are <0.1 µg/l and <0.03 µg/l, respectively. Immulite 2000 t-PSA assay from Diagnostics Product Corporation (Gwynedd, Wales) consists of a sequential non-competitive bead-based assay protocol with chemiluminescence detection. The detection limit is 0.02 µg/l. All samples and standards were run in duplicate and the mean value of both analyses was used for assay comparison. Statistical analysis was carried out with the aid of Analyse-it software, version 1.71 for Microsoft Excel.

3. Results

All the different components necessary for the construction of the PSA immunostrip had a key influence in both the performance and final sensitivity. We bore in mind that a t-PSA cut-off value of 4 µg/l has been accepted extensively and a 0.25 f/t-PSA ratio significantly improves the rate of prostate cancer detection in the t-PSA diagnostic grey zone of 4–10 µg/l. Therefore, a detection limit for both f-PSA and t-PSA of 1 µg/l or lower was an absolute requirement. This was achieved by carefully selecting the different materials used for the immunostrip construction as well as the experimental conditions for optimum strip performance. All the studies carried out are detailed below. These parameters were studied consecutively, one at a time, keeping other conditions constant.

1. Selection of membrane material. The choice of a suitable porous material for antibody immobilisation and sample flow was based on previous studies carried out by our research group. Nitrocellulose membranes were tested together with Porex polyethylene-based material. Porex membranes were found to be more robust and did not require a blocking step following antibody coating. In order to prevent conjugate non-specific binding, nitrocellulose membranes do require either this step or the addition of blocking agents to the sample pad in most applications. This normally has a negative effect on their capillary flow properties and the overall performance of the test (Verheijen et al., 1998).

2. Choice of antibodies to f-PSA and t-PSA and optimum concentration for membrane coating. Capture lines for f-PSA and t-PSA were produced using those monoclonal antibodies described in the experimental section. Three different antibodies to f-PSA that interact with epitopes located on different parts of this protein were tested. Under the same experimental conditions PSA19 mAb produced pink lines of higher colour intensity on the strip, which would render a better sensitivity in any resulting device. Likewise, the optimisation of both f-PSA and t-PSA concentrations for membrane coating should be carried out in order to obtain the required device sensitivity but, also, the same colour intensity for a particular f-PSA standard solution. Different antibody concentrations and ratios for anti-f PSA and anti-t PSA were assayed. A concentration of 2 g/l anti-f-PSA and 1 g/l anti-t-PSA were chosen accordingly. Working in this way, the resulting immunostrip test was easy to read and interpret since a simple visual observation could give us an estimate of how much the f-PSA and t-PSA content in a particular serum sample differed.
3. Optimisation of anti-PSA colloidal gold conjugate concentration for conjugate pad loading. Different volumes were assayed and 15 µl of the colloidal gold conjugate solution was the minimum amount that did not limit the sensitivity of the assay and which produced a reliable pink colour on all f-PSA, t-PSA and control lines.

4. Optimisation of sample volume. A volume of 80 µl provided an effective dissolution of the conjugate from the conjugate pad and the desired sensitivity for the detection of both PSA species without exceeding the absorbent capacity of the sink pad.

5. Washing step. The addition of a washing step to the assay performance allowed us to stop the immunoassay at a set incubation time of 15 min and thus minimise background readings thereby yielding better detection limits in the resulting device. As described in Section 5, a method of automatically carrying this out is proposed such that this washing step would not require further manipulation by the user after sample addition, which could negatively influence the reproducibility and reliability of the immunostrip.

Overall, the expected 1 µg/l detection limit was obtained for both PSA species, this being defined as the lowest PSA concentration that could be easily detected by the naked eye.

Fig. 2 depicts several strip tests carried out with f-PSA and PSA-ACT standards of different concentrations. The fact that PSA-ACT only bound to the anti-total PSA antibody capture line is an indication of the specificity of the PSA19 mAb towards f-PSA, i.e. this antibody shows negligible cross-reactivity with other PSA species. In addition, the colour intensity of the t-PSA capture line was similar at a fixed PSA concentration regardless of the PSA standard used (f-PSA or PSA-ACT). This demonstrates the equimolar activity of the PSA10 mAb towards both f-PSA and PSA-ACT. As pointed out in the experimental section, all the standard solutions used in this work were prepared in female serum. Female serum seldom contains detectable PSA levels so it is ideally suited to compensate for any matrix effects that could affect the performance of the device.

Dose–response curves were recorded for f-PSA and t-PSA using f-PSA standards over a concentration range of 1–100 µg/l and found to fit to a quadratic equation up to a concentration of 60 µg/l (Fig. 3). A calibration curve for t-PSA was also recorded with PSA-ACT standards and, as expected, its trend was nearly identical to the one obtained with f-PSA standards.

The reproducibility and stability of the device was tested for both f-PSA and t-PSA using a 10 µg/l f-PSA standard solution. The relative standard deviation
(RSD) values for within and day-to-day assay precision lay between 12% and 17% in all cases (n = 10 and n = 25, respectively). The latter was estimated by running two tests over 25 consecutive days. Stability tests were carried out by running two assays every 2 days during a period of 2 months. After obtaining the corresponding colour intensity values, stability values were estimated as percentage of the intensity values measured at day 0. All these were in the range of 85% to 115%, which lay between the day-to-day reproducibility values of the test, as stated above. Hence, it could be affirmed that the strips were stable for at least 2 months when stored at either room temperature or at 4 °C in a sealed plastic bag under desiccated conditions.

Fifty-one male serum samples were analysed with our device and two conventional immunoassays. The results were compared using both linear regression analysis and Bland–Altman plots (Bland and Altman, 1986, 1999). The correlation found between the different methods is shown in Table 1. Slope and Pearson coefficient values close to 1 indicated a strong linear correlation between the two methods. Hence, acceptable correlation for t-PSA among the different assays was found in all cases. f-PSA levels were estimated with the strip and the CanAg kit. The strip gave positive results in just 12 samples out of the 51 analysed. This meant that 39 serum samples had a PSA concentration below the f-PSA detection limit of the strip of 1 μg/l. Although the sample size appeared to be somewhat small for statistical analysis, the same studies were carried out and showed that the correlation found between the strip and the CanAg test was quite limited (Table 1).

According to Bland and Altman analysis (Bland and Altman, 1996), the Pearson correlation coefficient (r) measures the strength of a relation between two sets of values, not the agreement between them. Also, the slope values indicate how good the correlation between them is but do not give any information about their degree of agreement. This became quite evident in our study if the intercept values of the linear equations were considered. These were around +0.5 for the comparison studies between the strip and the two immunoassays, thereby indicating that even if the slope values approach 1, the strip test might underestimate both immunoassays.

The agreement between the assays could be easily assessed using Bland–Altman difference plots. These plots in Fig. 4 show the mean of the differences together with the 95% confidence intervals around the mean (95% CI) and the limits of agreement (LA). The estimated values for these parameters are shown in Table 2. The limits of agreement define the interval where 95% of the values are expected to lie. The mean value of the differences between two assays gives an estimation of the bias, i.e. the tendency for one method to exceed the other. A systematic bias appears when the 95% confidence intervals calculated around that mean do not include the 0 value. Comparison of strip and CanAg tests in Fig. 4A shows that just 2 samples out of 51 were outside the LA. Considering the t-PSA cut-off value of 4 μg/l, one of those samples was tested as positive with the strip and negative with the immunoassay whereas the other one was tested positive in both analyses. No apparent systematic bias existed between these two tests (95% CI = −0.14–0.50).

Fig. 4B and Table 2 show the plot of results of the comparison study between strip and Immulite assays. Here, 3/48 serum samples lay outside the LA but none were significant as they tested positive using either test. In contrast, a systematic bias was detected (95% CI = −1.02–(−0.26)). The strip test appeared to underestimate the Immulite test. It should also be pointed out that the LA and 95% CI were wider than those of the comparative study between strip and CanAg assays.

A similar study carried out between the CanAg and Immulite immunoassays (Fig. 4C and Table 2) showed significant differences in their performance. Here, 4/48 samples were outside the LA, but this interval was significantly narrower than in the previous two studies.

---

**Table 1**

<table>
<thead>
<tr>
<th>PSA species</th>
<th>PSA range (μg/l)</th>
<th>Assays</th>
<th>n</th>
<th>Slope</th>
<th>Intercept</th>
<th>Pearson correlation coefficient</th>
<th>Two-tailed P value&lt;sup&gt;b,c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-PSA</td>
<td>1–20</td>
<td>Strip-CanAg</td>
<td>51</td>
<td>0.85</td>
<td>0.50</td>
<td>0.87 ± 0.09</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Strip-Immulite</td>
<td>48</td>
<td>1.02</td>
<td>0.53</td>
<td>0.87 ± 0.09</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CanAg-Immulite</td>
<td>48</td>
<td>1.16</td>
<td>0.15</td>
<td>0.97 ± 0.02</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>f-PSA</td>
<td>1–3</td>
<td>Strip-CanAg</td>
<td>12</td>
<td>0.55</td>
<td>0.19</td>
<td>0.46 ± 0.36</td>
<td>0.1360</td>
</tr>
</tbody>
</table>

<sup>a</sup> ± values for slope, intercept and Pearson coefficient indicate the 95% confidence intervals around the calculated parameter.

<sup>b</sup> Student t-approximation.

<sup>c</sup> P value indicates the probability that the values given by two assays for one sample are not related.
Also, there was a clear bias, indicating that the CanAg test underestimated the Immulite test (95% CI = 1.05–0.62). This may be the source of the discrepancies found when those assays were individually compared with the immunostrip under development in this study (Fig. 4 A, B).

A first approach to demonstrate the accuracy of the test could be developed by taking those values given by the two conventional immunoassays as the real serum PSA levels. Thus, positive and negative test results for t-PSA were classified as either true positive or false positive and true negative or false negative by comparison with the t-PSA values obtained with the conventional immunoassays and considering the cut-off value of 4 μg/l (Table 3). Herein, sensitivity and specificity were defined as the capability of the test to discriminate between PSA levels higher and lower than 4 μg/l, respectively. It was shown that the strip yielded 20 true positive results out of 23 positives and 22 negative results out of 28 negatives given by the CanAg test. These data provide a sensitivity of 87% and a specificity of 79%. Despite the systematic bias found in the comparison with the Immulite kit, the values for these two parameters were similar to those obtained with the

### Table 2
Bland–Altman analysis carried out from values shown in graphs of Fig. 4

<table>
<thead>
<tr>
<th>PSA species</th>
<th>PSA range (μg/l)</th>
<th>Assays</th>
<th>n</th>
<th>Bias (95% CI)</th>
<th>Lower limit of agreement (95% CI)</th>
<th>Upper limit of agreement (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-PSA</td>
<td>1–20</td>
<td>Strip-CanAg</td>
<td>51</td>
<td>0.18 (0.14–0.50)</td>
<td>−2.05 (−2.58–(−1.52))</td>
<td>2.40 (1.87–2.94)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Strip-Immulite</td>
<td>48</td>
<td>−0.64 (−1.02–(−0.26))</td>
<td>−3.20 (−3.84–(−2.57))</td>
<td>1.93 (1.30–2.56)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CanAg-Immulite</td>
<td>48</td>
<td>−0.84 (−1.05–(−0.62))</td>
<td>−2.30 (−2.66–(−1.94))</td>
<td>0.63 (0.27–0.99)</td>
</tr>
<tr>
<td>f-PSA</td>
<td>1–3</td>
<td>Strip-CanAg</td>
<td>12</td>
<td>0.48 (0.12–0.83)</td>
<td>−0.63 (−1.17–(−0.08))</td>
<td>1.58 (1.04–2.12)</td>
</tr>
</tbody>
</table>

Fig. 4. Bland–Altman plots for (A) t-PSA immunostrip-CanAg, (B) t-PSA immunostrip-Immulite, (C) t-PSA CanAg-Immulite, and (D) f-PSA immunostrip-CanAg. Solid line—zero bias. Dashed lines: a—mean difference value (bias); b—95% confidence interval around the mean; c—limits of agreement (bias ± 1.96S.D.).
The greater differences in both cases were observed in samples with t-PSA values around the cut-off value of 4 μg/l. The f-PSA comparison study shown in Fig. 4D and Table 2 also demonstrated a systematic bias or overestimation of the strip compared with the CanAg test. All samples appeared to be within the LA. In spite of this limited agreement between the two assays, by calculating the sensitivity and specificity of the assay for a cut-off value of 1 μg/l, chosen on the basis of the f/t PSA ratio of 0.25, and considering the 51 samples under study, the test shows a sensitivity of 100%. Thus, the test predicted all those positive results obtained with the CanAg immunoassay. The specificity was 87% because 6 samples were wrongly predicted as positive.

Results with the strip and CanAg tests were also compared by considering an f-/t-PSA cut-off value of 0.25 for those samples with a t-PSA value in the diagnostic grey zone of 4–10 μg/l. A total of 28 serum samples lay within that range. The strip test yielded a sensitivity of 81% whereas the specificity was 29%.

### 4. Discussion

At present, PSA appears to be the only reliable tumour marker for the early detection of prostate cancer. In spite of the fact that a cut-off value of 4 μg/l for t-PSA in serum has largely been accepted, it is known that this criterion alone lacks specificity as other prostatic diseases induce an increase of the serum PSA levels. Significant improvement in the diagnostic accuracy have been demonstrated by considering the measurement of the ratio of f-PSA to t-PSA and establishing a cut-off value of 0.25 within the t-PSA range of 4–10 μg/l.

The anticipation of large screening programmes for prostate cancer has prompted the scientific community to look for alternative approaches other than conventional immunoassay platforms. These would have to be equally reliable but much simpler, faster and more suitable for decentralised testing. Lateral or flow through immunoassays on a strip is very attractive in this context. Indeed, during the last decade several PSA immunostrip tests have become commercially available. However, several studies carried out with some of these tests revealed pitfalls mainly related to lack of performance uniformity (Jung et al., 1999), limited sensitivity and specificity and end-user problematic readout (Oberpenning et al., 2003). While lack of performance uniformity could be attributable to the anti-PSA antibodies used, the other parameters seem to be an intrinsic problem of all immunostrp tests especially around the device detection limits. Nevertheless, t-PSA strip tests have been shown to be useful and may certainly contribute to the development of larger studies for the early detection of prostate cancer disease.

This work was carried out with the aim of demonstrating the feasible fabrication of an immunostrip test for the simultaneous detection of free and total PSA in serum. Using a suitable combination of monoclonal antibodies to f-PSA and t-PSA immobilised on a polyethylene-based membrane, a colloidal gold anti-PSA tracer and f-PSA standards prepared in female serum, the immunostrip developed provided reliable responses up to a PSA concentration of 60 μg/l and had limits of detection as low as 1 μg/l for both PSA species within a 15 min analysis time. A washing step was introduced in order to improve assay readout reliability especially at low PSA concentrations and in that case the overall assay time was less than 20 min.

Method comparison studies for t-PSA demonstrated a lack of assay uniformity between the two commercial immunoassays. Although there was a good assay correlation, the CanAg test seemed to underestimate the results given by the Immulite test. This different behaviour was clearly reflected in the comparative studies carried out with the immunostrip under development and both commercial immunoassays. While regression analysis yielded similar correlation data in both studies, Bland–Altman plots were significantly different. No apparent bias was found between the strip and the

<table>
<thead>
<tr>
<th>PSA species</th>
<th>PSA range (μg/l)</th>
<th>Immunoassay</th>
<th>n</th>
<th>Positive test results</th>
<th>Negative test results</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-PSA</td>
<td>1–20</td>
<td>CanAg</td>
<td>51</td>
<td>20 True positive</td>
<td>6 False positive</td>
<td>3 True negative</td>
<td>87 79</td>
</tr>
<tr>
<td></td>
<td>1–20</td>
<td>Immulite</td>
<td>48</td>
<td>23 True positive</td>
<td>3 False positive</td>
<td>15 True negative</td>
<td>77 83</td>
</tr>
<tr>
<td>f-PSA</td>
<td>0–3</td>
<td>CanAg</td>
<td>51</td>
<td>6 True positive</td>
<td>6 False positive</td>
<td>39 True negative</td>
<td>100 87</td>
</tr>
<tr>
<td>f-/t-PSA</td>
<td>4–10&lt;sup&gt;a&lt;/sup&gt;</td>
<td>CanAg</td>
<td>28</td>
<td>17 True positive</td>
<td>5 False positive</td>
<td>2 True negative</td>
<td>81 29</td>
</tr>
</tbody>
</table>

<sup>a</sup> t-PSA range-diagnostic grey zone.
CanAg assay. The calculated limits of agreement and confidence intervals were reasonable, thus indicating a fairly good agreement between the two methods for the detection of t-PSA. In contrast, the same studies carried out with the strip and Immulite assay clearly demonstrated that the t-PSA results given by the strip underestimated the Immulite t-PSA data. Likewise, the calculated limits of agreement and confidence intervals were wider than in the case of the comparison with the CanAg assay. This better agreement with the CanAg assay was expected, provided that the antibodies used for the construction of the strip were also supplied by CanAg Diagnostics and having observed the different performance of the two commercial immunoassays.

Comparison studies for f-PSA showed both lack of correlation and limited agreement between the CanAg test and the immunoassay. Although all samples that tested positive for f-PSA are within the calculated LA, this interval and the calculated 95% CI around the upper and lower LA were quite wide. Also, the strip test is likely to produce values for f-PSA slightly higher than the immunoassay. It should be pointed out that a group of 12 serum samples was rather small for statistical analysis and could be considered as not highly representative of the overall assay performance.

All samples tested with the immunoassay and the CanAg immunoassay were run in duplicate. Although rarely done, this is important in method comparison studies in order to give an estimation of the repeatability of each method for each individual sample and check whether this was likely to have a significant effect on their limited agreement (Bland and Altman, 1999). According to Bland and Altman, the repeatability of two methods being compared limits the amount of agreement that can be achieved. Comparing the values of the so-called repeatability coefficient, shown in Table 4, it is evident that the repeatability of the CanAg tests for both f-PSA and t-PSA is better than that of the immunoassay. This coefficient also defines an interval (± repeatability coefficient) where 95% of the differences between the two repeated values are expected to lie. Mathematically, this interval is equivalent to the interval defined by the LA in method comparison studies. Thus, as proposed by Bland and Altman, if this interval is similar to the estimated LA for the CanAg-strip comparison study, then the limited agreement observed between the two tests is explained by their lack of repeatability. In this context, t-PSA studies yielded acceptable repeatability measurements using both tests whereas the poor repeatability coefficient obtained with the strip for f-PSA resulted in a poor agreement with the CanAg test, as already discussed above.

The limited reproducibility of the immunoassay described herein is likely to arise mainly from the hand-coating of membrane and conjugate pad rather than any further manipulation by the user. A significant improvement in the reproducibility of this device could be envisaged by using automatic dispensers for spray-coating of the capture antibodies on the membrane and the colloidal gold conjugate on the pad, and also

---

**Table 4**
Repeatability coefficients and limits of agreement of the comparison studies between strip and CanAg test

<table>
<thead>
<tr>
<th>PSA species</th>
<th>Test</th>
<th>Repeatability coefficient</th>
<th>Limits of agreement b</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-PSA</td>
<td>Immunoassay</td>
<td>1.90</td>
<td>−2.05–2.40</td>
</tr>
<tr>
<td></td>
<td>CanAg</td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>f-PSA</td>
<td>Immunoassay</td>
<td>1.07</td>
<td>−0.63–1.58</td>
</tr>
<tr>
<td></td>
<td>CanAg</td>
<td>0.53</td>
<td></td>
</tr>
</tbody>
</table>

* a Repeatability coefficient = 1.96 × SD of the differences between the two repeated values.
* b Values taken from Table 2.
by inserting an automatic washing step within the operation of the immunostrip, as explained below.

Further analysis carried out with the aim of giving an estimation of the validity of the developed immunostrip yielded acceptable sensitivity and selectivity values for the measurement of f-PSA and t-PSA individually. It is noteworthy that in spite of the limited agreement found between the different tests for f-PSA, the immunostrip correctly detected 100% of the samples with an f-PSA concentration above the threshold value of 1 µg/l. The selectivity was also good (87%) though was affected by the trend of the immunostrip to exceed the f-PSA values given by the CanAg test. The latter also affects the selectivity obtained for the f-/t-PSA ratio (29%). The sensitivity of the immunostrip for the detection of f-/t-PSA ratio is, by contrast, maintained to those accepted levels obtained for the measurement of t-PSA alone (81%).

5. Future prospects

As pointed out above, the addition of a washing step during immunostrip performance is desirable in order to decrease background readings and thus improve the sensitivity of the assay. Fig. 5 depicts a means of inserting such a washing step within the strip device in order to avoid further handling and manipulation. Small storage blisters filled with a certain volume of the washing solution, to which a pad of a wicking material was stuck underneath, could be easily inserted in a plastic cassette also containing the strip. The cassette includes a lid that is provided with a window for the reading of the strip and a switch for the control of an articulated arm. Such arm exhibits a spike for the piercing of the blister. After the addition of sample and incubation for 15 min, the switch is pressed. This moves the arm that in turn induces the piercing of the blister and pushes it down to contact the strip. At the same time, the arm stops the sample flow by removing the sample and conjugate pads out of contact with the strip. Hence, the washing solution is allowed to flow, effectively removing all the unbound species from the membrane. This approach was assessed by carrying out preliminary repeatability studies and results were similar to those obtained above when the washing step was carried out manually.

Acknowledgements

Cesar Fernandez-Sanchez is grateful to the European Commission for the award of a Marie Curie Fellowship (Contract Number HPMF-CT-2002-01992) of the Human Potential programme. We thank Dr. Carmen Martin-Ruiz from the Department of Gerontology of the University of Newcastle upon Tyne for assistance with the colour intensity measurements.

References


Polascik, T.J., Oesterling, J.E., Partin, A.W., 1999. Prostate specific antigen: a decade of discovery—what we have learned and where we are going. J. Urol. 162, 293.


