

Some Analytical Considerations on the Measurement of Prostate-Specific Antigen

Bert G. Blijenberg¹, Ries Kranse², Ineke Eman¹ and Fritz H. Schröder²

¹ Department of Clinical Chemistry, University Hospital Rotterdam, Rotterdam, The Netherlands

² Department of Urology, University Hospital Rotterdam, Rotterdam, The Netherlands

Summary: In this study, various analytical aspects of the determination of serum prostate-specific antigen are described as applied to the Abbott IMx PSA and the Hybritech Tandem-E PSA assays.

We used mainly specimens from a prostate cancer screening study in progress. A very good comparability between the assays proved to exist in our hands. The long-run variation (16 months) was also rated as acceptable, both for the IMx and the Tandem-E method. The method of choice, Tandem-E, showed good reagent stability over this period.

We found, however, a difference in accuracy (Tandem-E \pm 8% higher values) that could not be explained by comparison with Tandem-R.

Introduction

In the second half of 1994 the Rotterdam section of the European Randomized Study of Screening for Prostate Cancer started (1). The goal is to evaluate the feasibility of a randomized screening study of prostate cancer locally and eventually, at the European level.

The criteria for biopsy of the prostate in the screening group are:

1. prostate-specific antigen concentration \geq 4.0 μ g/l
2. suspicious digital rectal examination
3. suspicious transrectal ultrasonography

In the definitive protocol the Hybritech Tandem[®]-E assay was chosen for the measurement of total prostate-specific antigen.

We have been using the Abbott IMx PSA assay for clinical purposes since early 1993. Both assays have been well-known in the field of clinical chemistry for a number of years. Nevertheless, we felt the need to assess some practical issues in our laboratory in relation to the screening protocol.

In this report we elaborate the comparability of both assays statistically as well as the stability of the Hybritech procedure.

Materials and Methods

Samples

All serum samples used were centrifuged within 2–3 hours after collection.

Most of the samples coming from participants of the screening study were analyzed either the same or the following day. In case of analysis on the following day they were stored at + 4 °C. The remaining screening samples came from the – 80 °C freezer. In the Results section the relevant information is presented.

The patient samples used were from our serum bank with well-characterized serum samples.

The quality control samples we used were either commercial samples (Bio Rad Lyphochek[®] Normal and Abnormal) or human serum pools. Regarding the serum pools, these were, after pooling divided in 0.3 ml aliquots and stored at – 80 °C.

Methods

The following methods were used exactly according to the instructions of the manufacturers:

1. Tandem[®]-E PSA, Hybritech Inc., (USA), a solid-phase two-site immunoenzymetric assay. Samples containing prostate-specific antigen are reacted with a plastic bead coated monoclonal antibody directed toward the prostate-specific antigen molecule, and with an enzyme-labelled monoclonal antibody directed against a distinctly different antigenic site on the same prostate-specific antigen molecule. After washing and incubation with enzyme substrate the amount of substrate turnover is measured colorimetrically.

2. Tandem[®]-R PSA, Hybritech Inc., (USA). The same measuring principle except for the last step, where the radioactivity is measured (¹²⁵I γ -radiation).

3. IMx[®]-PSA, Abbott Laboratories, (USA). Monoclonal antibody-coated microparticles capture the prostate-specific antigen analyte and are then reacted with a polyclonal goat antibody-alkaline immunoconjugate directed against prostate-specific antigen. After conversion of the substrate the fluorescent product, 4-methylumbelliferone, is measured.

In all cases single measurements were performed.

Statistical analysis

All statistical evaluations were done either by the regression method of *Passing & Bablok* (2) or by the differential procedure

as described by *Bland & Altman* (3, 4) or both. The latter is, for short, a test on the equality of analytical methods by assessment of bias and error. Here, we only applied the linear transformation of the results obtained with the methods under study.

Results

We started the study by comparing 122 specimens from patients with known prostate cancer at various stages. We only chose one detection criterium i. e. a reasonable division of the results over the whole measuring range 0–50 $\mu\text{g/l}$.

In figure 1 a graph is given of this comparison.

The comparison with cancer patients was followed by a comparison with fresh samples from the screening population. For practical reasons we had to perform analyses with the IMx on the day of the participants visit while the Tandem-E analysis was done on the next day.

Figure 2 shows the results after the transformation according to the method of *Bland & Altman*.

Later, we repeated this experiment with a smaller sample collective and with different IMx and Tandem-E rea-

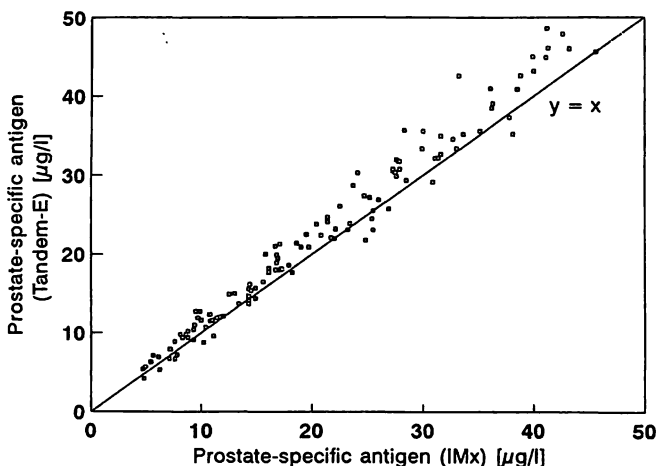


Fig. 1 Comparison of IMx assay with Tandem-E using 122 samples from patients with prostate cancer.

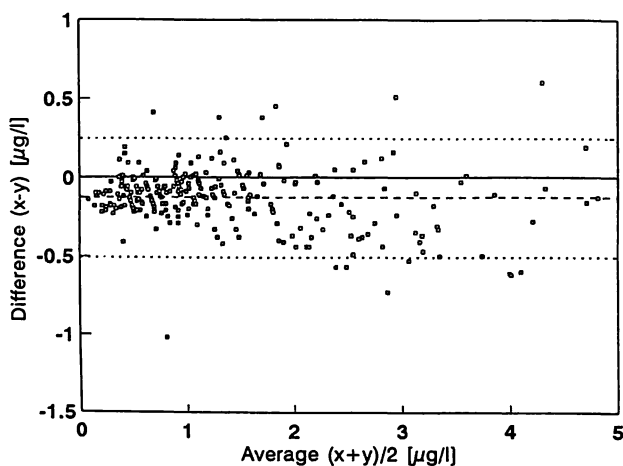


Fig. 2 Difference plot of 262 observations randomly sampled from a prostate cancer screening population. IMx(x): day 1, Tandem-E(y): day 2.

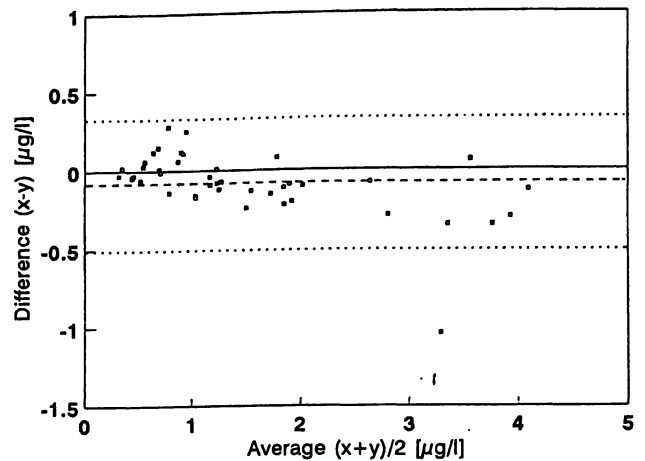


Fig. 3a Difference plot of 42 observations randomly sampled from a prostate cancer screening population. IMx(x): day 1, Tandem-E(y): day 2. Dashed lines: mean difference (---) and 95% confidence interval (···).

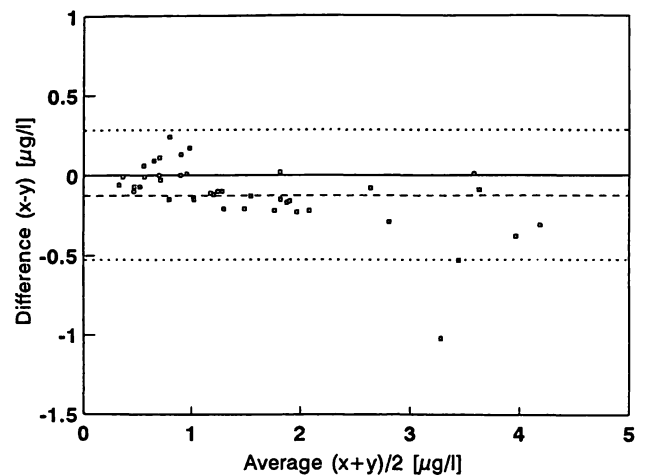


Fig. 3b The same samples as plotted in figure 3a. IMx(x): day 1, Tandem-E(y): day 1. Dashed lines: mean difference (---) and 95% confidence interval (···).

gent lots. We also included a simultaneous comparison of the IMx and Tandem-E assay on the day of visit as well as on the following day. In figure 3a + b the separate IMx and Tandem-E results are compared again with application of the *Bland & Altman* procedure.

Because of the differences found in all IMx and Tandem-E comparisons, we decided, not having standard prostate-specific antigen preparations, to compare the Tandem-E procedure with the Tandem-R assay. This was done on two separate occasions with different reagent lots, both for the Tandem-E and the Tandem-R assay. In this experiment we used freezer-stored samples, aliquotted for both assays and analyzed simultaneously. Both comparisons showed the same picture. Therefore, we combined the results in figure 4.

The regression lines of all previous experiments were calculated. In table 1 the relevant statistical information is given.

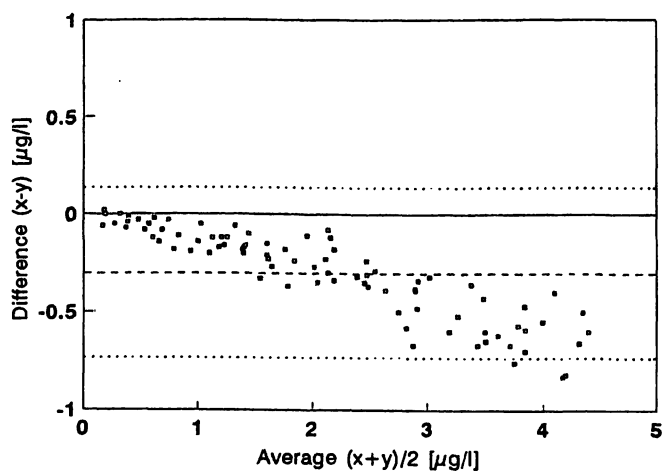


Fig. 4 Difference plot of 87 observations from freezer stored screening samples.

Tandem-R = x, Tandem-E = y.

Dashed lines: mean difference (---) and 95% confidence interval (···).

In addition, in table 2 all remaining information related to the examination of the differences of all paired methods, is given.

Finally we checked the imprecision of both assays during a long period. The day-to-day coefficients of variation are given in table 3. In figure 5 we made a subdivi-

sion of the various Tandem-E reagent lots we used in this 16-month period.

Discussion

From all statistical information given it is clear that we found excellent correlations between the various assays both for the patient and the screening samples.

However, we cannot deny that we detected an accuracy problem. Based on the average values of all specimen collectives and the data from the regression analysis we saw differences between the IMx and Tandem-E assays of, roughly speaking, 8%, the Tandem-E results being higher.

The difference plots, figures 2 and 3, show explicit evidence of systematic bias between both assays because in all comparisons the zero values are not included in the respective confidence intervals of the mean differences (tab. 2).

On the other hand, looking at the lines of agreement, we feel that the agreement is acceptable, judging clinically. There are very few outliers. Of course, there is a difference between the clinical and the screening samples in relation to medical decision making. Regarding this as-

Tab. 1 Statistical information on all method comparisons

Comparison	Prostate-specific antigen, range (µg/l)	Regression equation	r	n
1. IMx (= x), Tandem-E (= y)	0-50	$y = 1.09x - 0.06$	0.986	122
2. IMx (= x), Tandem-E (= y)	0-5.0	$y = 1.05x + 0.07$	0.983	262
3. IMx (= x), Tandem-E (= y)	0-5.0	$y = 1.11x - 0.06$	0.987	42
4. IMx (= x), Tandem-E (= y)	0-5.0	$y = 1.11x - 0.03$	0.988	42
5. IMx, IMx	0-5.0	$y = 0.99x - 0.02$	0.995	42
6. Tandem-E, Tandem-E	0-5.0	$y = 0.99x - 0.03$	0.997	42
7. Tandem-R (= x), Tandem-E (= y)	0-5.0	$y = 1.17x - 0.04$	0.997	87

Remarks

1. Comparison 1 = patient samples, comparison 2-7 = screening samples.

2. Comparisons 2, 3, 5 and 6: x-axis = day 1, y-axis = day 2, comparison 4 done on the same day with fresh samples and the same for 7 with stored sera.

Tab. 2 Statistical differences between all paired methods

Comparison	Prostate-specific antigen		Mean difference (x-y) [95% confidence interval](µg/l)
	Average		
	x (µg/l)	y (µg/l)	
1. IMx (= x), Tandem-E (= y)	20.6	22.3	-1.69 [(-2.06) - (-1.32)]
2. IMx (= x), Tandem-E (= y)	1.31	1.43	-0.13 [(-0.15) - (-0.11)]
3. IMx (= x), Tandem-E (= y)	1.48	1.56	-0.08 [(-0.04) - (-0.12)]
4. IMx (= x), Tandem-E (= y)	1.48	1.60	-0.12 [(-0.06) - (-0.18)]
5. IMx, IMx	1.48	1.42	0.06 [0.02 - 0.10]
6. Tandem-E, Tandem-E	1.60	1.56	0.03 [0.01 - 0.05]
7. Tandem-R (= x), Tandem-R (= y)	1.96	2.26	-0.30 [(-0.35) - (-0.25)]

Remarks

1. The same sequence of comparisons was used as described in table 1.

2. Average = mean of all results.

Tab. 3 Overall day-to-day imprecision IMx and Tandem-E

Control sample	Prostate-specific antigen, average ($\mu\text{g/l}$)	CV (%)	n
a. IMx			
Lyphochek Normal	1.8	5.6	302
Lyphochek Abnormal	5.5	3.5	302
b. Tandem-E			
Pool 1	4.9	3.6	302
Pool 2	1.7	10.3	125
Pool 3	2.4	4.1	169

Remarks

1. IMx reagent batches: 8.
2. Tandem-E reagent batches: see figure 5.

pect it is clear that the well accepted value of the upper limit of the prostate-specific antigen reference range, 4 $\mu\text{g/l}$, is at least questionable in relation to different methods.

Comparing the results shown in figure 3 with those in figure 2 the same picture arises. In addition, one may also conclude that the moment of measurement is not critical (figs. 3a + b). However, the statistical data mentioned in table 2 do point to a very small difference between measuring on the first or the second day. We will study this phenomenon later on in more detail.

In the comparison study with the screening samples, we did not split up the samples into the two categories benign prostate hyperplasia and prostate cancer because a great deal of the samples came from participants who were not studied any further. However, we assumed that most of them had benign prostate hyperplasia based on their age.

This aspect may be of value because it is known that the IMx assay is more sensitive for free prostate-specific antigen than Tandem-E which measures equimolarly (5).

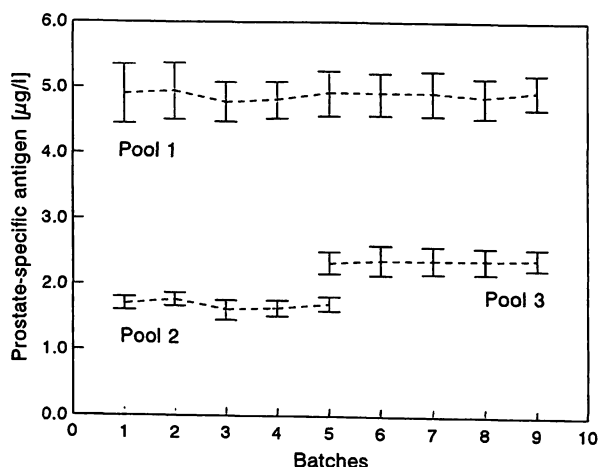


Fig. 5 Day-to-day variation of Tandem-E assay measured with 3 different human serum pools and 9 different reagent batches. All measurements per batch were between 20 and 40 days. Average values and 95% confidence intervals are plotted.

Our results are comparable to those published by *Brawer et al.* (6), *Dnistrian et al.* (16) and *Wright et al.* (7) though the study populations and practical conditions differ. *Brawer et al.* show small differences in assay behaviour by plotting IMx and Tandem-E results with benign prostate hyperplasia and prostate cancer patient samples in the prostate-specific antigen range of 2.0–10.0 $\mu\text{g/l}$. The number of specimens in both cases however, is limited.

We struggled with the results shown in figure 4, regarding the difference in accuracy between the Tandem-E and the Tandem-R assays. It was not our intention to check this methodological aspect. On the contrary, it proved to be a logical consequence of the studies described earlier (figs. 1–3), because very often the Tandem-R assay has the status of reference technique in the medical literature.

According to the information of the manufacturer both assays should show equality with application of the same antibodies. We accept that information, however, the detecting systems differ considerably. The situation may be comparable to what *Zucchelli et al.* showed in their comparison studies with 8 CA 19-9 assays, all with different labels (8). They found a large variation in results despite the use of the same (Centocor) antibodies.

Our findings concerning the range studied, i.e. up till 5.0 $\mu\text{g/l}$, were confirmed by *Hall et al.* (9) and by most of the quality assessment surveys held in The Netherlands and Germany.

At the moment we have no explanation for the various accuracy differences we found. The Tandem-E assay is a partly automated method that we ran exactly according to the manufacturer's instructions. The IMx is fully automated. The Tandem-R assay is the only manual assay and performed by us with calibrated glassware. The availability of internationally accepted prostate-specific antigen standards should help us solve this dilemma (10, 11).

The last item we want to describe concerns the imprecision of the prostate-specific antigen assays we compared. Recent literature discusses the variability of especially the IMx method (12, 13). Here, we do not want to go into detail in analyzing the performance of this method because it is clear that the overall coefficients of variation are acceptable as can be seen in table 3 and moreover, in line with what is published elsewhere (14, 15).

We prefer to focus on the Tandem-E assay. Table 3 also shows acceptable coefficients of variation for the Tandem-E method. The same holds for the various batches of reagents we used during a 16-month period, as can be seen in figure 5. We have no explanation for the small dip in the results obtained with pool 2.

In conclusion we may state that the Tandem-E PSA assay proved to be a fairly robust method in our hands, well applicable to our desired goal.

Acknowledgements

Many thanks are due to *J. Langstraat* for practical support and to *A. Copper-Stuamer* for secretarial help.

References

1. Schröder FH, Damhuis RAM, Kirkels WJ, De Koning HJ, Kranse R, Nijs HGT, Blijenberg BG. European randomized study of screening for prostate cancer – the Rotterdam pilot studies. *Int J Cancer* 1996; 65:145–51.
2. Passing H, Bablok W. A new biomedical procedure for testing the equality of measurements from two different analytical methods. *J Clin Chem Clin Biochem* 1983; 21:709–20.
3. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; i:307–10.
4. Hollis S. Analysis of method comparison studies. *Ann Clin Biochem* 1996; 33:1–4.
5. Zhou AM, Tewari PC, Bluestein BI, Coldwell GW. Multiple forms of prostate-specific antigen in serum: differences in immunorecognition by monoclonal and polyclonal assays. *Clin Chem* 1993; 39:2483–91.
6. Brawer MK, Daum P, Petteway, Wener MH. Assay variability in serum prostate-specific antigen determination. *The Prostate* 1995; 26:1–6.
7. Wright GL, Fritsche HA. An evaluation of four commercially available immunoassays for measuring prostate-specific antigen. Presented at XXII Meeting of International Society for Oncodevelopmental Biology and Medicine; 1994 Sept 18–22; Groningen (The Netherlands). *Tumor Biol* 1995; 16:317.
8. Zucchelli GC, Pilo A, Chiesa MR. Growing use of nonisotopic CA 19-9 immunoassays increases between-laboratory variability. *Clin Chem* 1993; 39:909–11.
9. Hall M, Johnson JT, Carr J. Assays for serodiagnosis of prostate cancer. *Lab Med* 1992; 23:607–12.
10. Stamey TA. Second Stanford conference on international standardization of prostate-specific antigen immuno-assays: September 1 and 2, 1994. *Urology* 1995; 45:173–84.
11. Murphy GP. The second Stanford conference on international standardization of prostate-specific antigen assays. *Cancer* 1995; 75:122–8.
12. Wener MH, Daum PR, Brawer MK. Variation in measurement of prostate-specific antigen: importance of method and lot variability. *Clin Chem* 1995; 41:1730–7.
13. Magic SE, Pennoyer JD. Between-lot/between-instrument variations of the Abbott IMx method for prostate-specific antigen. *Clin Chem* 1992; 38:2343.
14. Fritsche HA, Babaian RJ. Analytical performance goals for measuring prostate-specific antigen. *Clin Chem* 1993; 39:1525–9.
15. Van Dalen A. Quality control and standardization of tumour marker tests. *Tumor Biol* 1993; 14:131–5.
16. Dnistrian AN, Schwartz MK, Smith CA, Nisselbaum JS, Fair WR. Abbott IMx® evaluated for assay of prostate-specific antigen in serum. *Clin Chem* 1992; 38:2140–2.

Received April 9/July 19, 1996

Corresponding author: Dr. B. G. Blijenberg, Academic Hospital, Rotterdam-Dijkzigt Department of Clinical Chemistry, Dr. Molewaterplein 40, NL-3015 GD Rotterdam, The Netherlands

