Clinical Chemistry 48:8
1279–1287 (2002)

Multicenter Evaluation of an Artificial Neural Network to Increase the Prostate Cancer Detection Rate and Reduce Unnecessary Biopsies

Carsten Stephan,[1,2*] Henning Cammann,[3] Axel Semjonow,[4] Eleftherios P. Diamandis,[2] Leon F.A. Wymenga,[5] Michael Lein,[1] Pranav Sinha,[6] Stefan A. Loening,[1] and Klaus Jung[1]

Background: **The percentage of free prostate-specific antigen (%fPSA) has been shown to improve specificity for the diagnosis of prostate cancer (PCa) over total PSA (tPSA). A multicenter study was performed to evaluate the diagnostic value of a %fPSA-based artificial neural network (ANN) in men with tPSA concentrations between 2 and 20 μg/L for detecting patients with increased risk of a positive prostate biopsy for cancer.**
Methods: **We enrolled 1188 men from six different hospitals with PCa or benign prostates between 1996 and 2001. We used a newly developed ANN with input data of tPSA, %fPSA, patient age, prostate volume, and digital rectal examination (DRE) status to calculate the risk for the presence of PCa within different tPSA ranges (2–4, 4.1–10, 2–10, 10.1–20, and 2–20 μg/L) at the 90% and 95% specificity or sensitivity cutoffs, depending on the tPSA concentration. ROC analysis and cutoff calculations were used to estimate the diagnostic improvement of the ANN compared with %fPSA alone.**
Results: **In the low tPSA range (2–4 μg/L), the ANN detected 72% and 65% of cancers at specificities of 90% or 95%, respectively. At 4–10 μg/L tPSA, the ANN detected 90% and 95% of cancers with specificities of 62% and 41%, respectively. Use of the ANN with 2–10 μg/L tPSA enhanced the specificity of %fPSA by 20–22%, thus reducing the number of unnecessary biopsies.**
Conclusions: **Enhanced accuracy of PCa detection over that obtained using %fPSA alone can be achieved with a %fPSA-based ANN that also includes clinical information from DRE and prostate volume measurements.**

Prostate-specific antigen (PSA)[7] is the most valuable tool for early detection, staging, and monitoring of prostate cancer (PCa) *(1)*. Because PSA is almost organ specific, but not cancer specific, increased PSA concentrations are found in patients with PCa but are also observed in patients with benign prostatic diseases. The low specificity of total PSA (tPSA) for the detection of PCa has led to the development of various approaches to improve the specificity of PSA. Measurements of PSA density, PSA transition zone density, and PSA velocity or doubling time and the use of age- or race-specific reference ranges have been proposed to increase the cancer detection rate per biopsy performed *(1, 2)*. Measurements of the molecular forms of PSA have been shown to improve specificity over tPSA alone *(3, 4)*. Use of the percentage of free PSA (%fPSA) in the tPSA range 4–10 μg/L can eliminate ~20–25% of unnecessary biopsies *(5, 6)*. For tPSA values <4 μg/L, the use of %fPSA has been reported to increase specificity *(7–9)*, but other studies have not confirmed this finding *(10, 11)*. It has been reported that %fPSA possibly correlates with prostate volume *(5, 12–15)*, age *(12, 13, 16)*, and other clinical variables *(1, 17)*. Recently,

Departments of [1] Urology and [6] Laboratory Medicine and Pathobiochemistry, and [3] Institute for Medical Biometry, University Hospital Charité, Humboldt University, D-10098 Berlin, Germany.
[2] Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ontario, M5G 1X5 Canada.
[4] Department of Urology, Westfälische Wilhelms-University, D-48129 Münster, Germany.
[5] Department of Urology, Martini Hospital, NL-9700 Groningen, The Netherlands.
*Address correspondence to this author at: Department of Urology, University Hospital Charité, Humboldt University Berlin, Schumannstrasse 20/21, D-10098 Berlin, Germany. Fax 49-30-450-515904; e-mail carsten.stephan@charite.de.
Received February 8, 2002; accepted May 20, 2002.

[7] Nonstandard abbreviations: PSA, prostate-specific antigen; PCa, prostate cancer; tPSA, total PSA; %fPSA, percentage of free PSA; DRE, digital rectal examination; ANN, artificial neural network; TRUS, transrectal ultrasound; and AUC, area under the ROC curve.

%fPSA has been proposed as a primary decision tool for first-time biopsy in men with a nonsuspicious digital rectal examination (DRE) within the tPSA range 4–10 μg/L, as well as for lower PSA values *(18, 19)*.

Since 1998, different models of logistic regression *(20, 21)* and artificial neural networks (ANNs) *(22)* have been introduced to improve the PCa detection rate *(23)*. These models include clinically relevant data and can add substantial information for detecting PCa while avoiding unnecessary biopsies in patients with benign prostates *(22, 24–30)*.

ANNs can predict the outcome of prostate biopsy in an individual patient better than traditional statistics and are able to handle a greater number of variables with nonlinear relations than logistic regression *(31)*. In a study of 656 men with tPSA in the range 4–10 μg/L, Finne et al. *(22)* could avoid approximately one-third of unnecessary biopsies at 95% sensitivity for the detection of PCa. For the lower tPSA range (2.6–4 μg/L), Babaian et al. *(29)* used a combination of three different ANNs and could avoid 63.6% of unnecessary biopsies. Recently, we evaluated a %fPSA-based ANN in 859 patients (tPSA range, 2–20 μg/L) using the variables tPSA, %fPSA, age, prostate volume, and DRE status *(30)*. The aim of the present study was (*a*) to establish the performance of this ANN in a multicenter study and (*b*) to demonstrate the improved accuracy of ANNs for detection of PCa compared with tPSA or %fPSA in clinical practice at different cutoffs.

Materials and Methods

STUDY GROUPS AND SAMPLES

The database consisted of complete data from 1188 patients (age range, 40–89 years). All 1188 men were investigated between March 1996 and October 2001 at one of six hospitals: University Hospital Charité (Berlin; n = 859); Westfälische Wilhelms-University (Münster; n = 97); Martini Hospital (Groningen), Nij Smellinghe Hospital (Drachten), and University Hospital (Groningen; n = 148); and Mount Sinai Hospital (Toronto, Canada; n = 84). The indications for referral were increased PSA values, obstructive micturition symptoms, abnormal DRE, or biopsy-conformed PCa, which explains the high number of PCa patients. In consequence, this population does not represent a screening population. Only individuals with a complete data set on PSA, %fPSA, age, prostate volume, and DRE status were used in this study. No patient received antiandrogen treatment before blood sampling, and all patients had a tPSA value in the range 2–20 μg/L.

All 721 PCa patients (median age, 65 years; range, 40–86 years; median prostate volume, 30 cm^3; range, 10–114 cm^3) were diagnosed histopathologically by ultrasound-guided sextant or octant prostate biopsy. Five of the six centers exclusively performed sextant biopsy, whereas in one center (University Hospital Charité, Berlin), the biopsy strategy was changed in September 1999 from sextant biopsies (n = 517) to octant biopsies (n = 342) with inclusion of two more lateral cores (near the apex).

The 467 patients with benign prostates (median age, 65 years; range, 43–89 years; median prostate volume, 38 cm^3; range, 14–114 cm^3) were all histopathologically confirmed. The histologic diagnosis on the basis of tissue obtained only by transurethral resection of the prostate or open adenomectomy (n = 32) or transrectal ultrasound (TRUS)-guided sextant or octant biopsy (n = 435) was benign prostatic hyperplasia with or without chronic prostatitis or no histologic abnormality, but no evidence of prostatic intraepithelial neoplasia. Differentiation based on the number of biopsy sessions and biopsy history was not performed.

PROCEDURES

PSA determinations. Serum samples were collected before any prostate manipulation or at least 3–4 weeks after an earlier manipulation such as prostate biopsy, DRE, prostate massage, TRUS, or cystoscopy. The blood samples were centrifuged within 2–3 h after venipuncture, and the supernatants (serum) were analyzed on the same day or stored at −20 °C no longer than 48 h until analysis was performed. Various studies showed that this approach of sampling would not have a detrimental effect on fPSA stability *(32, 33)*.

In all six centers, total and free PSA were assayed using the IMMULITE PSA and IMMULITE Free PSA assays (Diagnostic Products). The assays are solid-phase, two-site, sequential chemiluminescent immunometric tests that are automatically performed on the IMMULITE automated analyzer with detection limits of 0.02 and 0.03 μg/L for fPSA and tPSA, respectively. The tests use both polyclonal and monoclonal antibodies specific for PSA or a monoclonal anti-PSA antibody specific only for fPSA. The analytical performance and comparisons with other PSA tests have been described previously *(34, 35)*.

Prostate volume and DRE. Only well-trained urologists in all participating centers determined prostate volume by TRUS. The COMBISON 330 was used in Berlin and Toronto, and the COMBISON 360 (Kretz Technik) was used in all three Holland centers and in Münster. The prolate ellipse formula [$\pi/6$(transverse diameter × anterior-posterior diameter × cephalocaudal diameter)], which is the most widely used in urology *(36)*, was applied as a sufficiently accurate method of prostate volume estimation. Corresponding to data obtained in other studies *(37)*, the intra- and interobserver variations in estimating the prostate volume by TRUS were between 12% and 15% at the University Hospital Charité, as described previously *(38)*, but they were not systematically studied in the other study centers. A DRE finding nonsuspicious for cancer was defined as negative and a finding suspicious for cancer as positive.

STATISTICAL ANALYSES

Conventional statistical calculations. We used the statistical software SPSS 10.0 for Windows and SigmaPlot 2001 for Windows (SPSS). The nonparametric Kruskal–Wallis test of variance, the Mann–Whitney *U*-test, logistic regression analysis with forward variable selection, and Spearman rank correlations (r_s) were carried out. The diagnostic validity of tPSA, %fPSA, and the ANN was evaluated by ROC curve analysis *(39)*. The areas under the ROC curves (AUCs) were compared by a nonparametric method using the software GraphROC 2.1 for Windows *(40)*. Significance was defined as $P \leq 0.01$.

ANN. The ANN models were constructed with the SPSS-extramodule Neural connection 2.0 (SPSS). The ANN was evaluated, trained, tested, and used to enhance the detection of PCa within the tPSA range 2–20 μg/L, including subsets of this range. We used a back-propagation network in which the input layer consisted of five variables: tPSA, %fPSA, patient age, prostate volume, and DRE status. The hidden layer used three neurons and the output layer had one neuron representing the output value as the probability of PCa. ANN comparisons including ROC curve analysis were studied with all 1188 patients as well as with subsets of trained and tested patient values within the different PSA ranges. For the whole PSA range, for example, we used data from 1000 patients for training, from which 100 patients were used for validation of the ANN within the training run. The established ANN was then tested using data from the remaining 188 patients.

For all five tPSA ranges, the ANN model was trained initially and then tested. The training and testing sessions were repeated 10 times so that each of the patients was set once in the testing group. This procedure of cross-validation avoids overtraining of the ANN. The sum of correctly identified patients in the 10 runs was the overall number of correctly classified patients. The developed back-propagation ANN is a feed-forward network. The activation function of the three hidden layers and the output layer was the hyperbolic tangent function, which produces output values between −1 and 1.

A Borland Delphi 5 computer program called "ProstataClass" was constructed for an input of the five variables and an outcome of a value between 0 and 1, which constitutes the cancer probability. With this program, it is possible to choose a 90% or 95% specificity cutoff for tPSA input values of 2–4 μg/L or a 90% or 95% sensitivity cutoff for tPSA input values of 4.1–10 and 10.1–20 μg/L. The output value automatically indicates the patient's risk depending on the used cutoff.

Results

CORRELATION OF %FPSA TO CLINICAL DATA AND ITS DIAGNOSTIC VALIDITY

The distribution for all 1188 patients, including their median tPSA values within the different tPSA ranges 2–4, 4.1–10, 10.1–20, 2–10, and 2–20 μg/L, is shown in Table 1. The tPSA was significantly lower for patients with benign prostates in all tPSA ranges ($P \leq 0.001$), but not in the tPSA range 10.1–20 μg/L ($P = 0.23$). %fPSA was significantly lower ($P \leq 0.001$) in PCa patients in all analyzed tPSA ranges.

In all 721 PCa patients, %fPSA was negatively correlated with tPSA ($r_s = -0.31$; $P \leq 0.01$). In PCa patients, %fPSA was positively correlated with age ($r_s = 0.28$; $P \leq 0.01$) and prostate volume ($r_s = 0.26$; $P \leq 0.01$).

In the 467 patients with benign prostates, %fPSA was positively correlated with age ($r_s = 0.29$; $P \leq 0.01$) and with prostate volume ($r_s = 0.33$; $P \leq 0.01$). All other variables, such as age, PSA, DRE, and ANN output, were negatively correlated with %fPSA. Age was not statistically different between patients with PCa or benign prostates across the different tPSA ranges.

Prostate volume was significantly higher in patients with benign prostates for all subgroups. Prostate volume as a factor influencing %fPSA was analyzed. The %fPSA was positively correlated with prostate volume in both PCa patients ($r_s = 0.26$; $P < 0.01$) and those with benign prostates ($r_s = 0.33$; $P < 0.01$) for the entire tPSA range. An abnormal DRE was reported in 551 (46%) of all patients, whereas the remaining 637 (54%) had nonsuspicious DRE results.

For all tPSA ranges (2–4, 4.1–10, 10.1–20, 2–10, and 2–20 μg/L), the specificities at 90% and 95% sensitivity and the sensitivities at 90% and 95% specificity for the absolute %fPSA values were calculated (Tables 2 and 3). These %fPSA values were the so-called "cutoffs" at specific points (e.g., 90% sensitivity or specificity). There was

Table 1. Patients investigated within the tPSA ranges 2–4, 4.1–10, 10.1–20, 2–10, and 2–20 μg/L.[a]

tPSA range, μg/L	All patients				PCa patients				Patients with benign prostates			
	n	tPSA, μg/L	% fPSA	Volume, cm³	n	tPSA, μg/L	% fPSA	Volume, cm³	n	tPSA μg/L	% fPSA	Volume, cm³
2–4	219	3.0	15.5	32	72	3.1	12.5	27	147	2.9[b]	16.2[b]	33[c]
4.1–10	606	6.7	11.4	33	374	6.8	9.8	29	232	6.5[b]	14.3[c]	39[c]
10.1–20	363	13.6	9.0	35	275	13.7	7.9	31	88	12.6	13.0[c]	55[c]
2–10	825	5.8	12.2	32	446	6.3	10.1	29	379	4.8[c]	15.5[c]	37[c]
2–20	1188	7.3	11.1	33	721	8.6	9.3	30	467	5.8[c]	15.0[c]	38[c]

[a] The median tPSA, %fPSA, and prostate volume values, as well as their significances between PCa patients and patients with benign prostates, are shown.
[b,c] Mann–Whitney *U*-test: [b] $P < 0.01$; [c] $P < 0.001$.

Table 2. Specificities at given sensitivities for tPSA, %fPSA, and the ANN value in all PSA ranges investigated.

tPSA range, μg/L	Sensitivity, %	Specificity,[a] %			Cutoffs[b]	
		tPSA	%fPSA	ANN	%fPSA	ANN
2–4	90	25 (19–32)	20 (15–27)	38 (31–45)	23.7	0.04
	95	19 (14–25)	14 (9.2–19)	22 (17–29)	26.8	−0.04
4.1–10	90	20 (16–25)	34 (29–40)	62 (57–67)	17.3	0.39
	95	10 (7.3–14)	21 (16–26)	41 (36–47)	19.4	0.27
10.1–20	90	12 (7.2–20)	40 (31–49)	62 (53–71)	15.1	0.45
	95	5.7 (2.3–12)	26 (19–35)	39 (30–48)	18.3	0.38
2–10	90	30 (26–34)	31 (27–35)	53 (49–58)	19.4	0.27
	95	21 (17–25)	13 (10–16)	33 (29–37)	24.1	0.03
2–20	90	31 (28–35)	34 (31–38)	61 (57–65)	17.6	0.24
	95	22 (19–25)	23 (2–26)	35 (32–39)	21.6	0.02

[a] Specificities and their 95% confidence intervals (in parentheses) at sensitivity cutoffs of 90% and 95% for all investigated tPSA ranges.
[b] Cutoffs for %fPSA and the values obtained with the ANN at the given sensitivities of 90% and 95%.

a clearly visible downward trend of %fPSA with higher tPSA values as shown in Tables 2 and 3.

ANN and ROC analyses

For the tPSA ranges 2–4, 4.1–10, 10.1–20, 2–10, and 2–20 μg/L, the ANN with the five variables tPSA, %fPSA, age, prostate volume, and DRE status was evaluated separately (Table 4). Comparison of ANN to logistic regression for all investigated PSA ranges showed no differences. The ANN and the logistic regression method correctly identified 80–82% and 79–81%, respectively, of all patients. However, this confirms our earlier findings and those of other studies *(30, 41)*. Table 2 summarizes the specificities for tPSA, %fPSA, and ANN for the given sensitivities at 90% and 95% for all investigated tPSA ranges. The number of unnecessary biopsies saved by use of %fPSA or ANN compared with tPSA was calculated as the differences between the specificity cutoffs. Table 3 shows the sensitivities for the given specificities 90% and 95%.

Use of the ANN demonstrated a significant further enhancement over %fPSA in avoiding unnecessary prostate biopsies, with specificities ~20–30% higher than those obtained for %fPSA alone. Compared with tPSA alone, the specificities in the different tPSA ranges increased even more. At given sensitivities, the ANN showed ~20–45% higher specificities.

The ANN generally gave an output value between 0 (low PCa risk) and 1 (high PCa risk). In some cases the value was <0 or >1, which was of no further relevance. On the basis of this ANN, we developed the computer program ProstataClass, in which a physician can easily input the five variables and obtain the ANN output.[8] One example is demonstrated in Fig. 1. In the PSA range 2–4 μg/L, Fig. 1 depicts a hypothetical patient 1 with a tPSA of 3 μg/L, a %fPSA of 8%, a prostate volume of 20 cm^3, and a nonsuspicious DRE, who should undergo a biopsy (Fig. 1A). A slight increase of the prostate volume gives no indication for a biopsy (Fig. 1B). If the age were lowered by 10 years, the same variables (prostate volume, 25 cm^3) would produce a biopsy recommendation. A 50-year-old hypothetical patient with tPSA of 5 μg/L, %fPSA of 10%,

[8] Copies of the program may be obtained from Dr. Stephan at carsten.stephan@charite.de.

Table 3. Sensitivities at given specificities for tPSA, %fPSA, and the ANN value in all PSA ranges investigated.

tPSA range, μg/L	Specificity, %	Sensitivity,[a] %			Cutoffs[b]	
		tPSA	%fPSA	ANN	%fPSA	ANN
2–4	90	6.9 (2.8–14)	28 (19–38)	72 (62–81)	9.6	0.32
	95	4.2 (1.2–11)	9.7 (4.7–18)	65 (55–75)	6.5	0.8
4.1–10	90	19 (15–22)	26 (23–30)	61 (57–66)	7.4	0.77
	95	7.5 (5.4–10)	13 (10–16)	48 (44–53)	5.6	0.83
10.1–20	90	5.5 (3.4–8.3)	44 (39–49)	73 (68–77)	7.2	0.72
	95	2.5 (1.2–4.8)	31 (26–36)	55 (50–60)	6.0	0.93
2–10	90	21 (18–24)	27 (23–30)	59 (56–64)	8.1	0.85
	95	14 (11–17)	14 (11–17)	35 (31–39)	6.0	0.89
2–20	90	26 (23–29)	32 (29–35)	64 (61–67)	7.5	0.77
	95	9.2 (7.5–11)	20 (18–23)	45 (42–48)	6.0	0.91

[a] Sensitivities and their 95% confidence intervals (in parentheses) at specificity cutoffs of 90% and 95% for all investigated tPSA ranges.
[b] Cutoffs for %fPSA and the values obtained with the ANN at the given specificities of 90% and 95%.

Table 4. PCa and benign prostatic hyperplasia patients correctly classified by ANN in the tPSA ranges 2–4, 4.1–10, 10.1–20, 2–10, and 2–20 μg/L.

	tPSA ranges, μg/L				
	2–4	4.1–10	10.1–20	2–10	2–20
No. of patients investigated	219	606	363	825	1188
No. of patients correctly classified	184	479	301	644	832
Percentage of patients correctly classified	84%	79%	83%	78%	78%

prostate volume of 25 cm^3, and a nonsuspicious DRE would show an ANN output of "risk", whereas the same patient with a %fPSA of 12% would have the output "no risk". With this program, it is possible to use the %fPSA-based ANN to decide whether to biopsy.

On the basis of multiple tests of the performance of the five individually established ANNs within the five various tPSA ranges, we also established ANNs for the PSA ranges 2–10 and 10.1–20 μg/L, according to the requirements necessary for urologic practice.

ROC analyses for tPSA, %fPSA, and the ANNs were performed for all tPSA ranges separately using the tPSA values, %fPSA values, and the calculated ANN values to draw the related ROC curves (Fig. 2, A–D). Because there were no differences between the curves calculated with matched data, as we recently recommended to avoid misdistribution of data *(39)*, the curves were calculated with all patients. Additionally, we compared the ROC curves for tPSA, %fPSA, and the ANN obtained in Berlin, as the center with highest number of patients (n = 859), with those obtained at the other centers (n = 329). There were no differences between the AUCs, showing that the overall outcome of the study previously published was not center affected *(30)*.

For all tPSA ranges, with the exception of 2–4 μg/L tPSA ($P = 0.33$), the ROC curves for %fPSA were significantly above those for tPSA ($P \leq 0.01$). All curves for the different ANNs were significantly above the curve for %fPSA (Fig. 2, A–D). In the tPSA range 2–4 μg/L, the ANN performed better than %fPSA at the specificity cutoffs of 90% and 95% compared with the 4–10 μg/L tPSA range (Table 3). Especially at the clinically important sensitivity cutoffs of 90% or 95% (PSA values >4 μg/L), the use of an ANN provided an advantage over use of %fPSA cutoffs alone.

Discussion

When PSA alone is used to predict the probability of PCa within the 4–10 μg/L tPSA range, ~75% of all biopsies will be negative. The discovery of molecular forms of PSA renewed clinical research in enhancing the specificity in this tPSA range and have been more beneficial than other calculated markers, such as PSA density and velocity or age-adjusted cutoffs *(1, 17)*. Within the last few years, it has been clearly demonstrated that the use of %fPSA can significantly improve specificity by ~15–25% compared with tPSA, with only a minimal loss of sensitivity in detecting PCa *(5, 6, 13)*. This has been shown for the 4–10 μg/L tPSA range as well as for the lower ranges, 2.6–4 and 2.5–4 μg/L tPSA *(7, 8)*. Our data confirm the improved diagnostic accuracy with %fPSA. Recommendations regarding %fPSA cutoffs have been discussed previously *(30)*.

ANNs were introduced into urologic decision making in 1994 by Snow et al. *(23)*. Since that time, several studies have been performed using ANNs or conventional algorithms to enhance the detection of PCa *(41, 42)*. In 1998, Carlson et al. *(20)* introduced a logistic regression model that included %fPSA, tPSA, and patient age. The authors found an 11% increase in specificity over the use of %fPSA

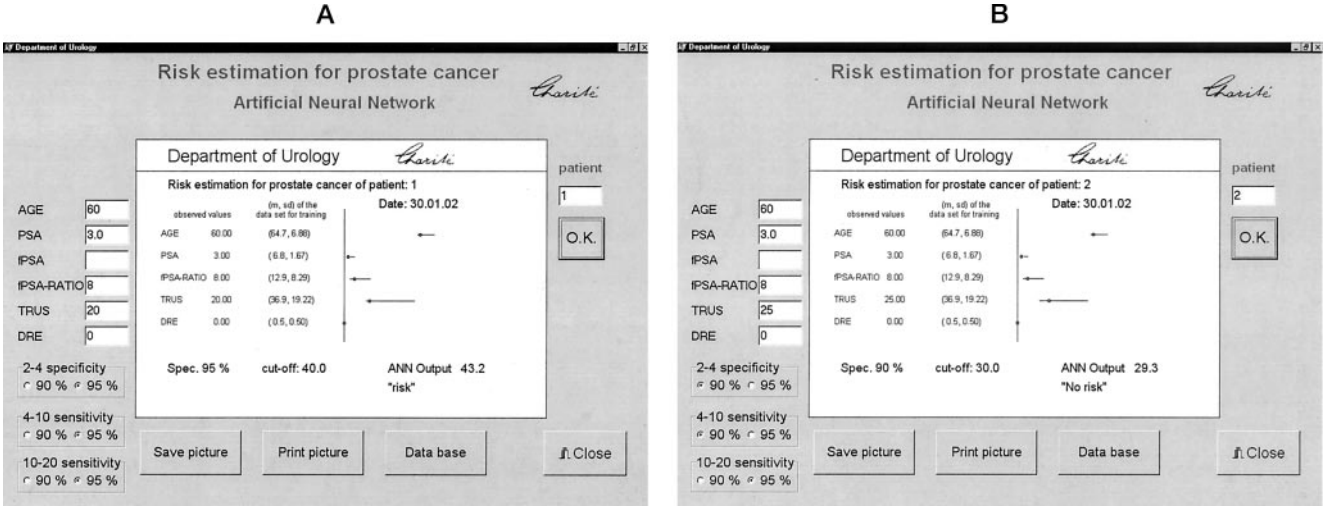

Fig. 1. Sample windows from the ProstataClass program.

Examples of the ANN output indicating risk (*A*) at the 95% specificity cutoff or no risk (*B*) at the 90% specificity cutoff based on individual factors keyed into the program for each hypothetical patient. Additional details are discussed in the text.

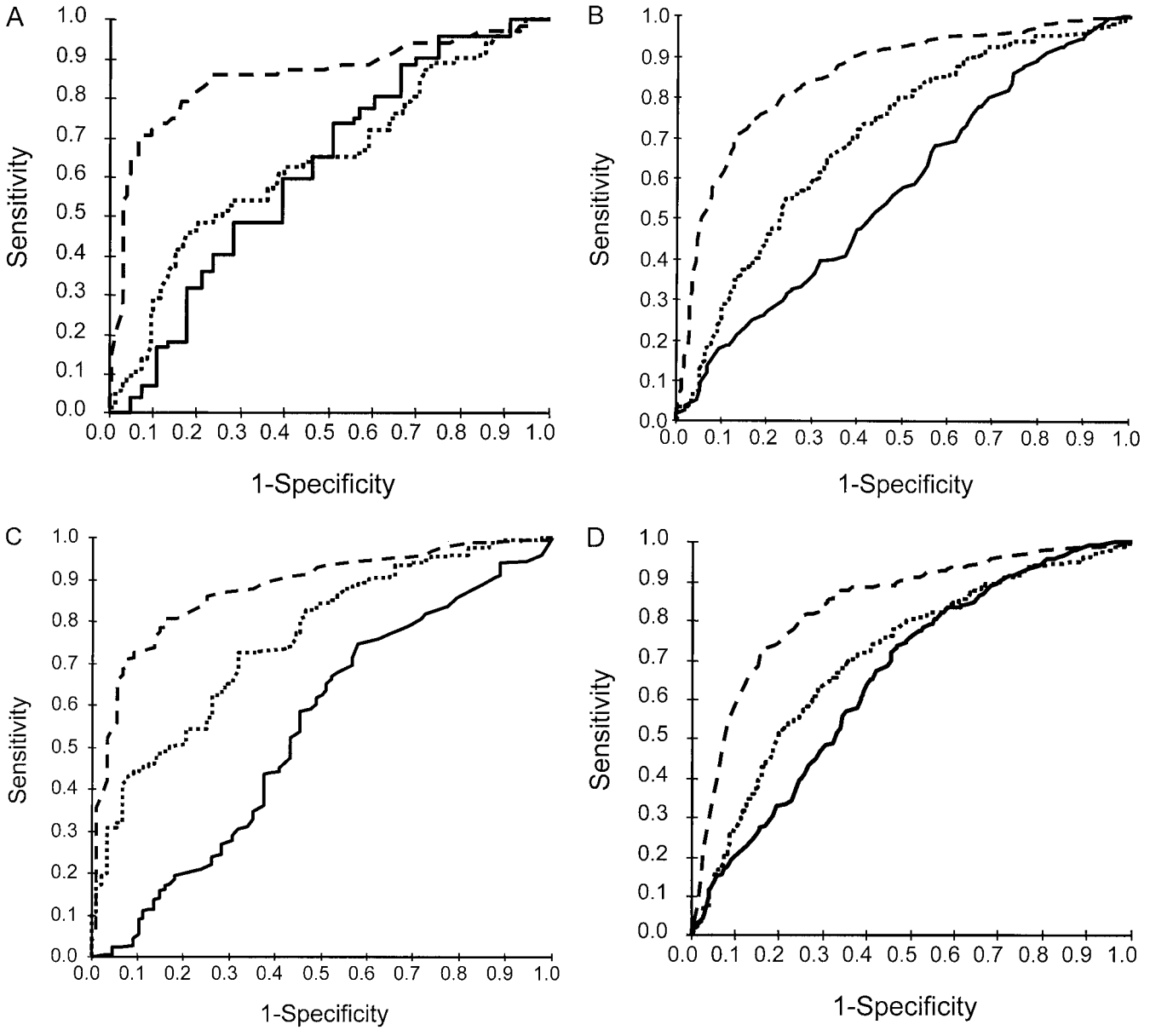

Fig. 2. ROC curves for tPSA (———), %fPSA (· · · ·), and the ANN (– – –).

Corresponding ROC curves for the tPSA ranges 2–4 μg/L (*A*), 4.1–10 μg/L (*B*), 10.1–20 μg/L (*C*), and 2–10 μg/L (*D*). Mean (SE) AUCs: (*A*), tPSA, 0.62 (0.04); %fPSA, 0.64 (0.04); ANN, 0.86 (0.03); (*B*), tPSA, 0.57 (0.02); %fPSA, 0.70 (0.02); ANN, 0.85 (0.02); (*C*), tPSA, 0.54 (0.04); %fPSA, 0.76 (0.03); ANN, 0.88 (0.02); (*D*), tPSA, 0.66 (0.02); %fPSA, 0.71 (0.02); ANN, 0.84 (0.01).

alone within the 4–20 μg/L tPSA range *(20)*. Virtanen et al. *(21)* used another logistic regression model and an ANN incorporating %fPSA, tPSA, DRE status, and heredity factors at tPSA from 3 to 10 μg/L. The results provided better diagnostic accuracy for PCa detection, with %fPSA and DRE status as the most powerful predictors. Data generated in the present study support DRE status and %fPSA as being the best input variables to detect PCa. A comparison of logistic regression and ANN models by Finne et al. *(22)* for the tPSA range 4–10 μg/L included prostate volume but not age. At 95% sensitivity, the specificities of %fPSA, logistic regression, and ANN were 19%, 24%, and 33% respectively.

The main aim of our study, however, was to establish a clinically usable program for the individual calculation of PCa risk. With differently calculated ANNs for the five tPSA ranges, we demonstrated significantly better performance for the ANNs compared with %fPSA in enhancing the specificity and sensitivity. ROC data and the comparative data for sensitivities and specificities, respectively, verify the high diagnostic validity and the advantage of using the ANN values instead of just %fPSA cutoffs for the decision to biopsy. The positive and negative predictive values were also significantly better for the ANNs compared with %fPSA or tPSA alone (data not shown). However, it should be pointed out that the ROC analysis

data and predictive values calculated from our referred population do not generalize to populations with other prevalences.

We recommend a first-time biopsy in the low tPSA range (2–4 μg/L) at a specificity of 95% to avoid general biopsies. In this low PSA range, the ANN considerably outperforms %fPSA and further enhances the sensitivity from 10% (%fPSA) to 65% (Table 3).

Regarding the high number of unnecessary biopsies, especially in the tPSA range 4.1–10 μg/L, we want to avoid a general biopsy within this tPSA range and recommend an ANN-based first-time biopsy based on the 90% or 95% sensitivity cutoffs. At these sensitivities, the specificity increased from 34% to 62% and from 28% to 72%, respectively.

For high tPSA values (10.1–20 μg/L), we recommend a general biopsy but would not rebiopsy if the ANN value indicates a cancer risk <5% (95% sensitivity). This strategy minimizes the risk of missing cancers and saves repeat biopsies in ~30% of all patients. Taken together, for all calculated tPSA ranges, the ANNs for 2–10 μg/L PSA and 10.1–20 μg/L PSA demonstrated significantly better performance than tPSA and %fPSA alone.

We are now routinely using the two ANNs for 2–10 and 10.1–20 μg/L tPSA. However, it must be emphasized that these data are valid only if the IMMULITE PSA and free PSA assays are used within this ANN because these were the assays used for the training and testing. Several studies have demonstrated large differences between PSA assays *(35, 43)*; therefore, the output data may change with other assays.

The most recent data on an ANN at the 2.5–4 μg/L tPSA range from Babaian et al. *(29)* demonstrated a remarkable enhancement of specificity at a 92% sensitivity cutoff, from 11% (%fPSA) to 62% (ANN). This ANN was established on a combination of three different ANNs using %fPSA, tPSA, age, and two other serum values as input variables, but not DRE or prostate volume *(29)*. Our data in the low tPSA range (2–4 μg/L) are analogous to these promising results. At 90% specificity, the ANN achieved 72% sensitivity, whereas tPSA (7%) and %fPSA (28%) had significantly lower results. The advantages of our ANN approach include its applicability for the whole tPSA range and the possibility of easy clinical use. The cutoffs designated for the different PSA ranges must be confirmed in further clinical use. A first-time biopsy based on %fPSA at 4–10 μg/L tPSA was recently proposed *(18)*. Our trained ANN may help avoid more unnecessary first-time prostate biopsies than the use of %fPSA alone. While this article was under review, one similar study was published *(44)*. The authors applied an ANN for the PSA range 2.5–10 μg/L with similar cutoffs for the low PSA range (95% specificity) and the 4–10 μg/L PSA range (95% sensitivity). Using PSA, %fPSA, prostate volume, and the additional volume-related ANN-input values, they achieved areas under the ROC curves for the ANN of 0.913 and 0.876 for the PSA ranges 2.5–4 and 4–10 μg/L, respectively. These results confirm the usefulness of PSA, %fPSA, and prostate volume-based ANNs for early detection of PCa.

In preliminary studies, other molecular forms of PSA and human glandular kallikrein 2 and other kallikreins have shown promise in improving the differentiation between PCa and noncancer patients *(45–49)*. Therefore, it is likely that in the near future these new molecular forms of PSA and other kallikreins (especially human glandular kallikrein 2) may add substantial information in extended ANNs including these markers and %fPSA.

As the ANN is used in clinical practice, more data will be generated that can be used to constantly update and improve the training of the algorithm. Undoubtedly, ANN technology can help in the diagnostic process for urologists, and available clinical and immunoassay data can be used to better assess the patient's risk of PCa and whether a biopsy is indicated.

In summary, we developed a diagnostic algorithm based on both immunoassay (Immulite PSA and IMMULITE free PSA) data and clinical data of age, prostate volume, and DRE status to enhance the performance of %fPSA to further reduce the number of unnecessary biopsies within the tPSA range 2–10 μg/L and to reduce the number of repeat biopsies at 10.1–20 μg/L tPSA. This ANN, which is based on data from referred populations, has to be validated prospectively in screening and other populations. In the future, ANNs could be used and hopefully expanded with new and promising input data to decrease the number of unnecessary biopsies in detecting PCa.

We gratefully acknowledge Paul Sibley for helpful suggestions. We thank Silke Klotzek, Sabine Becker, and Birgit Vogel for technical assistance. This work was supported by the following sources and grants: Merck Sharp & Dohme (to C.S.), Funds of the German Chemical Industry (No. 400770 to K.J.), and the SONNENFELD-Stiftung (to S.A.L.).

References

1. Polascik TJ, Oesterling JE, Partin AW. Prostate specific antigen: a decade of discovery–what we have learned and where we are going. J Urol 1999;162:293–306.
2. Nixon RG, Brawer MK. Enhancing the specificity of prostate-specific antigen (PSA): an overview of PSA density, velocity and age-specific reference ranges. Br J Urol 1997;79(Suppl 1):61–7.
3. Lilja H, Christensson A, Dahlen U, Matikainen MT, Nilsson O, Pettersson K, et al. Prostate-specific antigen in serum occurs predominantly in complex with α_1-antichymotrypsin. Clin Chem 1991;37:1618–25.
4. Stenman UH, Leinonen J, Alfthan H, Rannikko S, Tuhkanen K, Alfthan O. A complex between prostate-specific antigen and α_1-antichymotrypsin is the major form of prostate-specific antigen in serum of patients with prostatic cancer: assay of the complex improves clinical sensitivity for cancer. Cancer Res 1991;51:222–6.
5. Catalona WJ, Smith DS, Wolfert RL, Wang TJ, Rittenhouse HG,

Ratliff TL, et al. Evaluation of percentage of free serum prostate-specific antigen to improve specificity of prostate cancer screening. JAMA 1995;274:1214–20.
6. Catalona WJ, Partin AW, Slawin KM, Brawer MK, Flanigan RC, Patel A, et al. Use of the percentage of free prostate-specific antigen to enhance differentiation of prostate cancer from benign prostatic disease: a prospective multicenter clinical trial. JAMA 1998;279: 1542–7.
7. Catalona WJ, Partin AW, Finlay JA, Chan DW, Rittenhouse HG, Wolfert RL, et al. Use of percentage of free prostate-specific antigen to identify men at high risk of prostate cancer when PSA levels are 2.51 to 4 ng/mL and digital rectal examination is not suspicious for prostate cancer: an alternative model. Urology 1999;54:220–4.
8. Catalona WJ, Roehl KA, Antenor JAV. Robustness of free PSA measurements to reduce unnecessary biopsies in the 2.6 to 4.0 ng/ml total PSA range. J Urol 2001;165(Suppl):282–3.
9. Jung K, Stephan C, Elgeti U, Lein M, Brux B, Kristiansen G, et al. Molecular forms of prostate-specific antigen in serum with concentrations of total prostate-specific antigen <4 μg/l—are they useful tools for early detection and screening of prostate cancer? Int J Cancer 2001;93:759–65.
10. Carlson GD, Calvanese CB, Childs SJ. The appropriate lower limit for the percent free prostate-specific antigen reflex range. Urology 1998;52:450–4.
11. Hammerer P, Graefen M, Henke RP, Haese A, Huland E, Huland H. Ratio free/total PSA (%f-PSA) in men with total PSA 0.5–3 ng/ml—improvement for prostate cancer detection? [Abstract]. J Urol 2000;163(Suppl):277A.
12. Mettlin C, Chesley AE, Murphy GP, Bartsch G, Toi A, Bahnson R, et al. Association of free PSA percent, total PSA, age, and gland volume in the detection of prostate cancer. Prostate 1999;39: 153–8.
13. Partin AW, Catalona WJ, Southwick PC, Subong EN, Gasior GH, Chan DW. Analysis of percent free prostate-specific antigen (PSA) for prostate cancer detection: influence of total PSA, prostate volume, and age. Urology 1996;48:55–61.
14. Haese A, Graefen M, Noldus J, Hammerer P, Huland E, Huland H. Prostatic volume and ratio of free-to-total prostate specific antigen in patients with prostatic cancer or benign prostatic hyperplasia. J Urol 1997;158:2188–92.
15. Stephan C, Lein M, Jung K, Schnorr D, Loening SA. The influence of prostate volume on the ratio of free to total prostate specific antigen in serum of patients with prostate carcinoma and benign prostate hyperplasia. Cancer 1997;79:104–9.
16. Lein M, Koenig F, Jung K, McGovern FJ, Skates SJ, Schnorr D, et al. The percentage of free prostate specific antigen is an age-independent tumour marker for prostate cancer: establishment of reference ranges in a large population of healthy men. Br J Urol 1998;82:231–6.
17. Stephan C, Jung K, Lein M, Sinha P, Schnorr D, Loening SA. Molecular forms of prostate-specific antigen and human kallikrein 2 as promising tools for early diagnosis of prostate cancer. Cancer Epidemiol Biomarkers Prev 2000;9:1133–47.
18. Lee CT, Scardino PT. Percent free prostate-specific antigen for first-time prostate biopsy. Urology 2001;57:594–8.
19. Stephan C, Lein M, Jung K, Schnorr D, Loening SA. Can prostate specific antigen derivatives reduce the frequency of unnecessary prostate biopsies? [Letter]. J Urol 1997;157:1371.
20. Carlson GD, Calvanese CB, Partin AW. An algorithm combining age, total prostate-specific antigen (PSA), and percent free PSA to predict prostate cancer: results on 4298 cases. Urology 1998; 52:455–61.
21. Virtanen A, Gomari M, Kranse R, Stenman UH. Estimation of prostate cancer probability by logistic regression: free and total prostate-specific antigen, digital rectal examination, and heredity are significant variables. Clin Chem 1999;45:987–94.
22. Finne P, Finne R, Auvinen A, Juusela H, Aro J, Maattanen L, et al. Predicting the outcome of prostate biopsy in screen-positive men by a multilayer perceptron network. Urology 2000;56:418–22.
23. Snow PB, Smith DS, Catalona WJ. Artificial neural networks in the diagnosis and prognosis of prostate cancer: a pilot study. J Urol 1994;152:1923–6.
24. Horninger W, Bartsch G, Snow PB, Brandt JM, Partin AW. The problem of cutoff levels in a screened population. Cancer 2001; 91:1667–72.
25. Montie JE, Wei JT. Artificial neural networks for prostate carcinoma risk assessment. Cancer 2001;91:1647–52.
26. Han M, Snow PB, Epstein JI, Chan TY, Jones KA, Walsh PC, et al. A neural network predicts progression for men with Gleason score 3+4 versus 4+3 tumors after radical prostatectomy. Urology 2000;56:994–9.
27. Han M, Snow PB, Brandt JM, Partin AW. Evaluation of artificial neural networks for the prediction of pathologic stage in prostate carcinoma. Cancer 2001;91:1661–6.
28. Ziada AM, Lisle TC, Snow PB, Levine RF, Miller G, Crawford ED. Impact of different variables on the outcome of patients with clinically confined prostate carcinoma. Cancer 2001;91:1653–60.
29. Babaian RJ, Fritsche H, Ayala A, Bhadkamkar V, Johnston DA, Naccarato W, et al. Performance of a neural network in detecting prostate cancer in the prostate-specific antigen reflex range of 2.5 to 4.0 ng/mL. Urology 2000;56:1000–6.
30. Stephan C, Jung K, Cammann H, Vogel B, Brux B, Kristiansen G, et al. An artificial neural network considerably improves the diagnostic power of percent free prostate specific antigen in prostate cancer diagnosis—results of a five year prospective investigation. Int J Cancer 2002;99:466–73.
31. Partin AW, Murphy GP, Brawer MK. Report on Prostate Cancer Tumor Marker Workshop 1999. Cancer 2000;88:955–63.
32. Arcangeli CG, Smith DS, Ratliff TL, Catalona WJ. Stability of serum total and free prostate specific antigen under varying storage intervals and temperatures. J Urol 1997;158:2182–7.
33. Woodrum D, French C, Shamel LB. Stability of free prostate-specific antigen in serum samples under a variety of sample collection and sample storage conditions. Urology 1996;48:33–9.
34. Jung K, Stephan C, Lein M, Henke W, Schnorr D, Brux B, et al. Analytical performance and clinical validity of two free prostate-specific antigen assays compared. Clin Chem 1996;42:1026–33.
35. Semjonow A, Brandt B, Oberpenning F, Roth S, Hertle L. Discordance of assay methods creates pitfalls for the interpretation of prostate-specific antigen values. Prostate Suppl 1996;7:3–16.
36. Terris MK, Stamey TA. Determination of prostate volume by transrectal ultrasound. J Urol 1991;145:984–7.
37. Bates TS, Reynand JM, Peters TJ, Gingell JC. Determination of prostatic volume with transrectal ultrasound: a study of intra-observer and interobserver variation. J Urol 1996;155:1299–300.
38. Meyer A, Jung K, Lein M, Rudolph B, Schnorr D, Loening SA. Factors influencing the ratio of free to total prostate-specific antigen in serum. Int J Cancer 1997;74:630–6.
39. Jung K, Stephan C, Lein M, Brux B, Sinha P, Schnorr D, et al. Receiver-operating characteristic as a tool for evaluating the diagnostic performance of prostate-specific antigen and its molecular forms—what has to be considered? Prostate 2001;46: 307–10.
40. Kairisto V, Poola A. Software for illustrative presentation of basic clinical characteristics of laboratory tests—GraphROC for Windows. Scand J Clin Lab Invest 1995;55(Suppl 222):43–60.

41. Sargent DJ. Comparison of artificial neural networks with other statistical approaches. Cancer 2001;91:1636–42.
42. Wei JT, Zhang Z, Barnhill SD, Madyastha KR, Zhang H, Oesterling JE. Understanding artificial neural networks and exploring their potential applications for the practicing urologist. Urology 1998; 52:161–72.
43. Fox MP, Reilly AA, Schneider E. Effect of the ratio of free to total prostate-specific antigen on interassay variability in proficiency test samples. Clin Chem 1999;45:1181–9.
44. Djavan B, Remzi M, Zlotta A, Seitz C, Snow P, Marberger M. Novel artificial neural network for early detection of prostate cancer. J Clin Oncol 2002;20:921–9.
45. Marks LS, Llanes AS, Linton HJ, Gasior CL, Millar LS, Mikolajczyk SD, et al. BPSA is a potential serum marker for benign prostatic hyperplasia (BPH) [Abstract]. J Urol 2001;165(Suppl):266.
46. Mikolajczyk SD, Millar LS, Marker KM, Rittenhouse HG, Wolfert RL, Marks LS, et al. Identification of a novel complex between human kallikrein 2 and protease inhibitor-6 in prostate cancer tissue. Cancer Res 1999;59:3927–30.
47. Mikolajczyk SD, Marker KM, Millar LS, Kumar A, Saedi MS, Payne JK, et al. A truncated precursor form of prostate-specific antigen is a more specific serum marker of prostate cancer. Cancer Res 2001;61:6958–63.
48. Nam RK, Diamandis EP, Toi A, Trachtenberg J, Magklara A, Scorilas A, et al. Serum human glandular kallikrein-2 protease levels predict the presence of prostate cancer among men with elevated prostate-specific antigen. J Clin Oncol 2000;18:1036–42.
49. Partin AW, Catalona WJ, Finlay JA, Darte C, Tindall DJ, Young CY, et al. Use of human glandular kallikrein 2 for the detection of prostate cancer: preliminary analysis. Urology 1999;54:839–45.